An Interdisciplinary Study of the Timbre of the Classical Guitar.

Caroline Traube

Music Technology
Department of Theory
Faculty of Music
McGill University
Montreal, Canada

October 2004

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Doctor of Philosophy.

2004/10/08

Abstract

This dissertation proposes an interdisciplinary approach for the study of the timbre of the classical guitar. We start by identifying the static control parameters of timbre, relating to the structural components of the guitar and the dynamic control parameters of timbre, relating to the gestures applied by the performer on the instrument. From the plucked string physical model (obtained from the tranverse wave equation), we derive a digital signal interpretation of the plucking effect which is a comb filtering. Then we investigate how subjective characteristics of sound, like timbre, are related to gesture parameters. The starting point for exploration is an inventory of verbal descriptors commonly used by professional musicians to describe the brightness, the colour, the shape and the texture of the sounds they produce on their instruments. An explanation for the voice-like nature of guitar tones is proposed based on the observation that the maxima of the comb-filter-shaped magnitude spectrum of guitar tones are located at frequencies similar to the formant frequencies of a subset of identifiable vowels. These analogies at the spectral level might account for the origin of some timbre descriptors such as open, oval, round, thin, closed, nasal and hollow, that seem to refer to phonetic gestures. In a experiment conducted to confirm these analogies, participants were asked to associate a consonant to the attack and a vowel to the decay of guitar tones. The results of this study support the idea that some perceptual dimensions of the guitar timbre space can be borrowed from phonetics. Finally, we address the problem of the indirect acquisition of instrumental gesture parameters. Pursuing previous research on the estimation of the plucking position from a recording, we propose a new estimation method based on an iterative weighted least-square algorithm, starting from a first approximation derived from a variant of the autocorrelation function of the signal.

Résumé

L'objet de cette thèse est une étude interdisciplinaire du timbre de la guitare classique. Dans un premier temps, nous identifions les paramètres statiques de contrôle du timbre, liés aux composantes structurelles de la guitare, et les paramètres dynamiques de contrôle du timbre, liés au geste instrumental exécuté par le guitariste sur son instrument. À partir du modèle physique d'une corde pincée (déduit de l'équation d'onde transversale), nous dérivons une interprétation numérique de l'effet de point de pinçage, qui est un filtrage en peigne. Ensuite, cette recherche explore la manière dont des attributs subjectifs du son, tels que le timbre, sont reliés à des paramètres du geste. Le point de départ de cette exploration est un inventaire de descripteurs de timbre couramment employés par des musiciens professionnels lorsqu'ils décrivent la brillance, la tonalité, la forme et la texture des sons qu'ils produisent sur leur instrument. Nous proposons une explication du caractère vocal des sons de guitares. Cette explication est fondée sur le fait que les maxima de la structure de filtre en peigne du spectre d'amplitude des sons de guitare sont situés à des fréquences similaires aux fréquences centrales des formants de certaines voyelles bien identifiables. Ces analogies spectrales entre sons de guitares et sons vocaux pourraient expliquer l'origine de certains descripteurs de timbre, tels que ouvert, ovale, rond, mince, fermé, nasal et creux, qui semblent faire allusion à des gestes phonétiques. Dans une expérience menée dans le but de confirmer ces analogies, des participants devaient associer une consonne à l'attaque et une voyelle à la partie harmonique de sons de guitare. Les résultats de cette étude soutiennent l'idée selon laquelle certaines dimensions du timbre de la guitare peuvent être empruntées de la phonétique. Finalement, nous nous intéressons au problème de l'acquisition indirecte de paramètres du geste instrumental. Poursuivant des recherches antérieures sur l'estimation de la position du point de pinçage à partir d'un enregistrement, nous proposons une nouvelle méthode fondée sur un algorithme récursif en moindres carrés pondérés, partant d'une première approximation déduite d'une variante de la fonction d'autocorrélation du signal.

Acknowledgments

I would like to extend my deepest thanks to my advisor Prof. Philippe Depalle for welcoming me in the Sound Processing and Control Laboratory, allowing me to pursue my PhD thesis in a stimulating and supportive setting. I thank him for his intellectual and moral support and for helping me regain confidence in my research project. I greatly admire his sense of ethics, his honesty, as well as the respect and generosity he shows towards his students.

I also would like to thank Prof. Marcelo Wanderley for his inextinguishable enthusiasm. His seminar on the gestural control of music was a genuine enlightenment in the course of my studies and had a definite impact on the content of my thesis.

I would also like to thank my former advisor Prof. Julius Smith at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University, where I began my doctoral research. To him I owe the initial idea of this project (the problem of estimating plucking position from a recording). My two-year stay at CCRMA was a uniquely enriching intellectual and cultural experience.

This research would not have been possible without the active collaboration of Peter McCutcheon, a great guitarist and head of the Guitar Performance Area at Université de Montréal. I am most appreciative of his gracious availability for discussions about timbre and plucking techniques of the classical guitar, and for long and mundane recording sessions. I also thank him for his friendship and for sharing his passion for birds.

During the past four years, I have been most fortunate to teach at Faculty of Music at Université de Montréal. I am grateful towards my colleagues, composers Robert Normandeau and Jean Piché, for their trust and friendship, and for giving me the opportunity to learn so much through teaching and interacting with music students.

I would like to thank Alain Gauthier, Jean-François Dessureault and Olivier Bélanger, the teaching assistants who helped me so efficiently with the different courses I taught.

I am also grateful for the respect and the enthusiasm students at Université de Montréal demonstrated towards my teaching and my research. I was fortunate to meet exceptional students who combine their talents and interests for music and science, and from whom I learned very much. In particular, I would like to thank Madeleine Bellemare, for helping review the language with so much care and attention. Through the interest she showed towards my research, she helped me immensely. I would also like to thank Alexandre Savard,

who insisted that I met his guitar teacher, Prof. Peter McCutcheon, whose contribution to my research was key. Alexandre assisted with the segmentation of guitar tones (a task through which he claims to have acquired absolute pitch for the B and D strings). He also helped me to verify various theories on the perception of guitar timbre through lengthy discussions.

My thanks go to the students in the Guitar Performance programme at Université de Montréal who participated in the study on timbre descriptors.

Thanks to the creators of LaTeX2e and its packages for this revolutionary text-processing system that prevents graduate students from pulling out their hair as they write their dissertations. And thank you to Vincent Verfaille for his friendship, who reassured and advised me through his own recent experience with theisis writing, and who helped me make the best use of LaTeX2e.

The completion of this thesis would not have been possible without the support of friends who were present during difficult times. My heartfelt thanks go to Regina Nuzzo for her true friendship, for always lending a listening ear and for imparting the most caring words; to Alain Beauregard for being a caring guardian angel; and to Isabelle Panneton, for her precious friendship and support.

Finally, I would like to thank my family and friends in Belgium, who were understanding and supportive, and who did everything in their power to help, despite the distance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

[...] to my mind, any community of musicological practice which excludes from consideration living musicians and restricts itself to accounts of frozen results of musical action, fails to be an inspiring community of inquiry about music.

Otto Laske [161] (p. 85).

Contents

1.1 The timbre of the classical guitar

1.1.1 The guitar as a miniature orchestra

The modern six-string guitar stems from sixteenth-century Spanish vihuela, which is rooted in antiquity. Throughout its history, it has nonetheless been treated as a second-class instrument, mostly due to its poor dynamic range. The recognition of the guitar as a concert instrument occured largely in the 19th century. Fernando Sor (1778-1839) was first of a long line of Spanish virtuosos and composers for the guitar.

Composers such as Hector Berlioz, Ludwig van Beethoven and Johannes Brahms valued the instrument's timbral qualities. Hector Berlioz, renowned for his great mastery of orchestral timbre, taught guitar in Paris for some years; in fact, it was one of the few instruments at which he was truly proficient.

The guitar is known as a "miniature orchestra", not only because it can sustain melody and accompaniment simultaneously or play polyphony like the piano, but also because of the vast array of timbral variations of which it is capable. The notion of the guitar as a small

orchestra has been reinforced by reviews of several guitar concerts in which critics praise performers for their ability to imitate the oboe, the violin, the harp, the trombone, the trumpet, the horn, and other orchestral instruments. 19th-century guitarists intuitively mimicked some distinctive aspects of the orchestral instruments' timbre. For example, Fernando Sor obtained an oboe-effect by plucking the string vertically to the soundboard with the nail very close the bridge. This does not emulate the attack of the oboe, but the spectrum produced by this method does indeed resemble the nasal tone of the oboe, at least in comparison with a usual guitar tone [30].

Another allusion to the orchestral guitar is by the father of the modern guitar, Francisco Tarrega (1852-1909), via his pupil Pascual Roch's *A Modern Method for the Guitar: School of Tarrega*. In the section entitled "Artistic and Beautiful Effects on the Guitar," Roch describes harp-tones, bell-tones, side-drum effects, bass-drum effects, trombone effects, and the clarinet or oboe effects and their production [28].

1.1.2 The voice of the guitar

The vocal quality of the guitar timbre has been noted many times. The early-romantic composer Franz Schubert is known to have played the instrument each morning and to have written many of his lieder at the guitar [30]. The guitar was for him particularly effective at evoking sung melodies. In his book on the school of Tarrega [28], Roch included a section explaining how to imitate the "Cracked Voice of an Old Man or Woman", sobbing, a stammerer, and a stammerer singing. The Russian historian Makaroff described a Spanish guitarist with very evocative terms: "The vibrato, when performed by Ciebra, was really divine – his guitar actually sobbed, wailed and sighed." [21]. Furthemore, guitarists often use words related to speech to describe their playing techniques. As Duncan states: "Articulation pauses before notes allow control of color and of rhythmic placement. They enhance the clarity of one's musical enunciation by providing space for notes to breathe" [23] (p. 62).

1.1.3 Timbre and musical expression

Timbre plays a major role in musical expression. However, musical expression has been traditionally related to expressive timing and dynamic deviations in performance [71]. Less attention has been given to how musical expression relates to timbre. This is probably

due to the difficulty of defining the features of timbre, which are related to the physical aspects of sound in very complex ways. On the other hand, pitch, duration and volume are perceptual phenomena that have fairly simple physical correlates.

1.2 Looking into a timbre subspace

1.2.1 From a macroscopic to a microscopic point of view

Timbre can be studied at different levels. From a *macroscopic* point of view, one may examine the differences between the timbre of a violin and the timbre of a guitar. From a *microscopic* point of view, one may examine the differences within these instrumental categories, such as subtleties between a Stradivarius and a Guarnerius violin, or a Ramirez and a Rubio guitar. Furthermore timbre can be examined from the performer's point of view, by analysing, for example, the difference between a note played ponticello (close to the bridge) and tasto (close to the nut) on the same instrument. This is the perspective we propose in this thesis.

1.2.2 Relationship between gesture and timbre

When examining timbre microscopically, the paramount importance of the performer is suddenly brought forth. From where does the sound truly originate? The instrument or the performer?

When investigating the timbre of a musical instrument, it is crucial to take into account the performer's actions, which are responsible for all the timbre variations attainable on an instrument. The object of the study is not the instrument alone but the interactive coupled system made of the performer and the instrument.

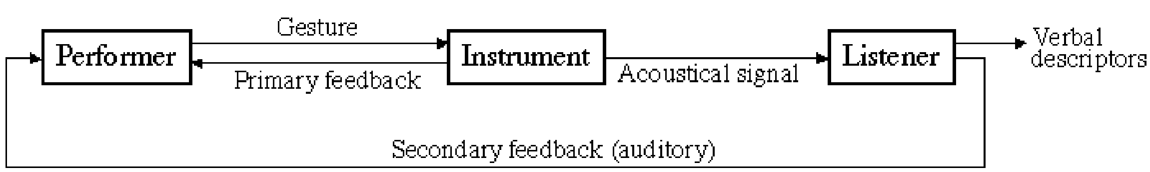


Fig. 1.1 The performance process loop.

Fig. 1.1 schematizes the exchange of information between the three elements of a performance process: the performer, the instrument and the listener. A musician is at the

same time a performer and a listener. The performer applies a gesture to the instrument, which in turn reacts to the gesture by producing a sound and by providing the performer with primary feedback, which can be visual, auditory (clarinet key noise, for instance) and tactile-kinesthetic [75], as well as with secondary feedback, which is auditory and corresponds to the sound produced by the instrument as perceived by the musician *listening* who can react to this information, as a musician *performing* by adjusting his/her playing techniques. The listener perceives the sounds produced by the instrument and attaches label to them. Expert performers/listeners are generally able to discriminate and intuitively describe a large variety of sounds produced by their instruments.

1.2.3 The verbal description of timbre: an oral tradition

On the guitar, different plucking techniques involve varying instrumental gesture parameters such as the finger position along the string, the inclination between the finger and the string (in a plane parallel to the string), the inclination between the hand and the string (in a plane perpendicular to the string), the degree of relaxation of the plucking finger, the choice of fingering on the neck of the guitar (string/fret combination), etc.

Among these parameters, the plucking position has the greatest effect on timbre. If the plucking point is closer to the bridge, the sound is <u>brighter</u>, <u>sharper</u>, <u>more percussive</u>. If the plucking point is closer to the middle of the string or the soundhole, the resulting sound is <u>warmer</u>, <u>mellower</u>, <u>duller</u>, as expressed by expert performers/listeners. This intuitive correlation between plucking position (a gesture parameter) and brightness (a perceptual dimension of timbre) is well-known and acknowledged by most guitarists. But it only summarily describes the timbral palette of the instrument.

Guitarists perceive subtle variations of instrumental gesture parameters and they have developed a very rich vocabulary to describe the brightness, the colour, the shape and the texture of the sounds they produce on their instruments. <u>Dark</u>, <u>bright</u>, <u>chocolatey</u>, <u>transparent</u>, <u>muddy</u>, <u>wooly</u>, <u>glassy</u>, <u>buttery</u>, and <u>metallic</u> are just a few of those adjectives.

The meaning of this often metaphorical vocabulary is transmitted from teacher to student, as an oral tradition. A very small number of guitarists (and performers in general) write about this vocabulary, which is so often taken for granted.

In the Western world, a standard notation for timbre never developed. In the East, however, a highly elaborate system of notation evolved for the timbres of the Ch'in, an

ancient Chinese seven-string lute. One of the earliest written accounts of this notation system is the *Sixteen Rules for the Tones of the Lute* by Leng Ch'ien (14$^{\text{th}}$ century B.C.E.). It describes in 150 to 200 special characters the techniques for performing the sixteen archetypical "touches" or tones of the lute, the names of which include "The Gliding Touch", "The Crisp Touch", "The Empty Touch" and "The Profound Touch." [25].

1.3 Questions and answers

Here are the questions that launched this research on the timbre of the classical guitar:

- What is the effect on sound of plucking parameters such as the plucking position?
- As gesture parameters are clearly perceived and recognized by experienced performers, is it possible to automatically extract parameters such as the plucking position from the analysis of a digital recording?
- How are different instrumental gestures related to different timbres on the guitar?
- How do guitarists control, perceive and verbally describe the timbre of their instruments?
- What is the acoustical basis of this vocabulary for the description of timbre?
- In particular, what is a "round sound"? What is "round" about a guitar sound?
- What is the "voice" of a guitar? Where does that vocal quality come from?

The answers to these questions appeared to lie at the intersection of many spheres of theoretical and practical knowledge and the need for interdisciplinarity imposed itself naturally, bridging across disciplines such as acoustics, signal processing, linguistics, psychology, music performance and pedagogy. The sources are accounts of research on topics as diverse as guitar acoustics, guitar playing techniques, timbre perception, speech production and perception and singing techniques. The answers to these questions do not only lie in books nor in the computer analysis and simulation of guitar tones. The seed to many answers sprouted from a fruitful collaboration with living musicians. Through questionnaires and interviews, we unearthed the practical knowledge and understanding of sound that

performers develop through years of practice, a knowledge that has been shared almost exclusively within the context of teaching the instrument practice.

This work would not have been possible without the collaboration of guitarists who enthusiastically agreed to patiently communicate their art to someone with absolutely no prior knowledge in their field. Studying the guitar in isolation would have been very limiting since the guitar does not play itself. The guitarist, as an agile-fingered puppeteer, enlivens an inanimate sounding object. The guitarist speaks and sings through the instrument; indeed the guitar is an extension of the guitarist's voice.

1.4 Contents and organization of this thesis

This thesis is divided into three parts that reflect each of the directions in which the research evolved. The first part examines the production of guitar tones while the second studies their perception. The third is devoted to the extraction of gesture and timbre parameters from a recording.

The first part is divided in four chapters. Chapter 2 presents all the structural components of guitar that may affect the timbre produced by the instrument. Since the sound of the guitar is determined not solely by the construction of its body, but also by the interaction between the player's fingers and the string, Chapter 3 covers the interaction between the strings and the guitarist's fingers. Chapter 4 describes the physical behaviour of the plucked string. The magnitude spectrum coefficients of an ideal plucked string are derived. Differences between an ideal string and a real string are presented. In Chapter 5, we present the digital signal processing interpretation of the plucking string physical model which is a comb filter. Then, the notion of "comb filter formant" is introduced. We also describe the digital modeling of plucked strings for waveguide-based synthesis, using a comb filter to simulate the localized plucking excitation and we explain how the comb filter delay should be set for a realistic reproduction of the performance.

The second part of the thesis, which concerns the perception of guitar tones, begins in Chapter 6 with a review of the main theories of timbre perception and the methods used to study the perception and the description of timbre. Chapter 7 contains an inventory of adjectives for the description of the timbre of the classical guitar with their subjective definitions and corresponding plucking techniques. Information about the vocabulary used

by guitarists to describe timbre was collected in two ways: from written questionnaires submitted to professional guitarists and from interviews with professional guitarists. In Chapter 8, the "phonetic mode" of timbre perception is introduced. The voice and the guitar are compared from different points of view. The way in which linguists and musicians describe the timbre of speech sounds is reviewed. The interesting fact is that there exists a large set of qualifying adjectives used for the description of guitar tones and speech sounds. Chapter 9 reports an experiment that was conducted in order to verify the perceptual analogies between guitar sounds and vocal sounds, based on the analogies that were found at the spectral level. In the experiment, participants were asked to associate a consonant to the attack and a vowel to the release of guitar tones. Chapter 10 presents all the parallels that can be drawn between phonemes – the elementary units of speech – and *sonemes* – the elementary units of instrumental music.

The last and third part concerns the indirect acquisition of instrumental gesture parameters. Chapter 11 describes signal processing techniques for the extraction of instrument gesture parameters specific to guitar playing such as the plucking location along the string.

Fig. 1.2 Symbolic picture illustrating a finger technique for the Ch'in, an ancient Chinese seven-string lute (from a Japanese manuscript copy of the *Yang-ch'un-t'ang-ch'in-pu*). 'The flying dragon grasping its way through the clouds' suggests that the touch should be broad and firm, the hand having more or less a clawing posture [25].

Part I

Guitar Timbre Production

Chapter 2

The Classical Guitar

People are captivated by the sound of the guitar, lured by its intimate voice – a voice not always warm and seductive, but by turns cool and clear, dry and witty, even angry and violent.
John Taylor [32] (p. 5).

Contents

A guitar sound is mainly determined by the construction of the guitar body, the material and dimensions of the string, the interaction between the strings and the guitarist fingers and the room acoustics. This chapter presents the structural components of classical guitar, which are static control parameters of the instrument's timbre. The dynamic control parameters, relating to the interaction between the strings and the guitarist's fingers, will be discussed in the next chapter.

2.1 General description of the classical guitar

2.1.1 Component parts of the guitar

From a descriptive point of view, the guitar can be broken down into several component parts: string, soundboard, and soundbox. The classical guitar has a thin, responsive soundboard and is strung with six nylon strings. The three treble strings are made of monofilament or multifilament nylon; until the 1940's, they were made of twisted sheep gut. The three bass strings consist of wire wrapped around a core of multifilament nylon; traditionally, this core was made of silk threads [30].

From a mechanical point view, the guitar consists of two coupled vibrators, the string and the body. The vibrating string, as it moves, alternately compresses and rarefies the surrounding air. Alone, it is not a good sound radiator because of its small dimensions when compared to the wavelength of the generated sound. In order to better radiate the sound, the string is connected through a bridge to a body acting as an impedance adapter. Although it is not strictly speaking an amplification as there is no increase in the total energy supplied to the instrument, the effect of the body is perceived as an amplification of the sound.

The important parts of the guitar body are the bridge and top plate, the ribs, the back plate, the air cavity and the soundhole.

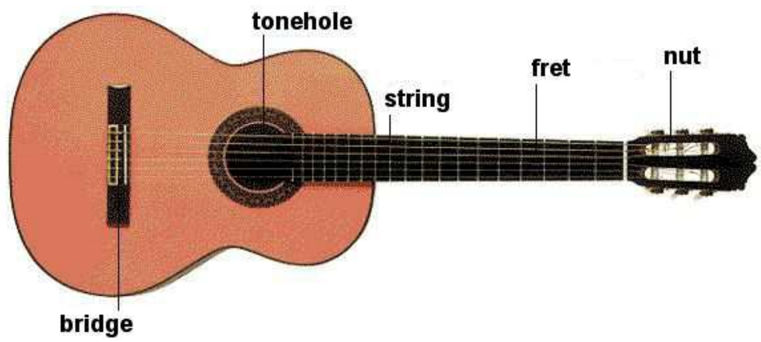


Fig. 2.1 Structural elements of the classical guitar.

The vibrating string applies a force on the bridge and pushes the top plate into vibration. The movement of the top plate sets into vibration the ribs, the air cavity, and the back plate. The sound wave radiated by the guitar body then travels from the instrument to the ears of the guitarist and of the audience.

2.1.2 Coupling between strings through the bridge

Table 2.1 gives the standard tuning for the six strings of a classical guitar.

String number	Note name	Standard tuning frequency (f_o)
6	E (Mi$_2$)	83 Hz
5	A (La$_2$)	110 Hz
4	D (Ré$_3$)	146 Hz
3	G (Sol$_3$)	202 Hz
2	B (Si$_3$)	248 Hz
1	E (Mi$_4$)	330 Hz

Table 2.1 Standard tuning for the six strings of a classical guitar.

When a string which shares the same pitch as (or has a common harmonic with) another string is plucked, the plucked string excites the unplucked string through sympathetic vibration and creates a tone which differs greatly from the normal plucked-string sound [30].

2.1.3 Fret rule for guitars

To set definite pitch relations between notes, metal inserts called frets are inset in a fretboard on the neck of guitars. The raised edges of the frets provide fixed lengths of string when the string is held down against them with a finger. The interval between successive frets is normally one equally tempered semitone. Guitar makers use a rule of thumb consisting in placing the frets *one-eighteenth the remaining length of the string* [6]. A string of length 17/18 of its original length is sharper by an interval of 98.9 cents, which is slightly less than an equally tempered semitone of 100 cents.

2.2 The body of an acoustic guitar

2.2.1 The top plate or soundboard

Since a thin string is not very efficient at moving air, it is necessary to connect the string to a soundboard whose greater surface area is more efficient at radiating vibrations. The link between the strings and the soundboard is the bridge. Not only holding the strings, the bridge determines the sound of the instrument by affecting how much of the string vibration is transmitted to the soundboard. Depending on its stiffness, the soundboard can be considered as a membrane or as a plate, and can simultaneously vibrate in a number of simple and complex modes [30]. It must be stiff enough to resist the tension from the strings so that the instrument will not bend; it also has to be light and flexible to respond well to the string vibrations.

The wood which is normally used for the top plate is spruce or cedar. Each wooden plate is unique in terms of its physical properties, which differ along and across the grain, vary from region to region on the plane, and depend on the way the panel is cut from the tree.

On the underside of the top plate, strips of wood called struts are glued in a pattern. They support the top plate against the string tension. The shape of the top plate modes and their contribution to radiation depends strongly on the chosen bracing pattern [15].

2.2.2 Coupling between string and soundboard

When an ideal string is set into motion between two completely stationary bridges, the only energy loss is due to friction within the string and friction between the string and

the surrounding air. But when one end of a string is coupled to a resonator, such as the soundbox of a guitar, energy is exchanged between the two systems.

The direction in which the string moves will determine the motion of the soundboard, but the soundboard's flexibility will also determine the movement of the string. The two systems affect each other (i.e. they are coupled) and the player can control the amount and quality of the force applied to the soundboard by the manner in which the string is plucked.

The string vibrational modes can couple with those of the body more or less strongly depending on the quality factor of the body modes. If the coupling is strong, the string and body modes are both perturbed so strongly that two totally new resonant modes of the string-body system appear instead of the uncoupled string and body modes. The strong coupling splits the resonant frequencies of the normal modes symmetrically about the unperturbed resonant frequencies, and both modes appear with the same damping. String and bridge move in phase at the lower frequency mode and in opposite phase at the higher frequency mode.

The plate will radiate most of the energy at its resonant frequencies very efficiently, but some of the energy at those frequencies will be fed back into the original vibrating string, as well as into the other five strings, through the movement of the bridge. If one or more of the unplucked strings resonate sympathetically with the driven string, the sound will be enhanced; if not, the energy loss will simply hasten the decay of the plucked string.

2.2.3 The guitar body as a Helmholtz resonator

An important resonant mode in the guitar body is due to the air resonance resulting from a standing wave created within the soundbox.

A Helmholtz resonator or Helmholtz oscillator is a container of gas (usually air) with an open hole (or neck or port). As illustrated on Fig. 2.2, a mechanical analog system is a spring (corresponding to the air volume) connected to a mass (corresponding to the opening). The resonant frequency of a Helmholtz resonator is inversely proportional to the square root of the volume of the body and is approximated with the following formula:

$$f = \frac{c}{2\pi}\sqrt{\frac{A}{VL}} \tag{2.1}$$

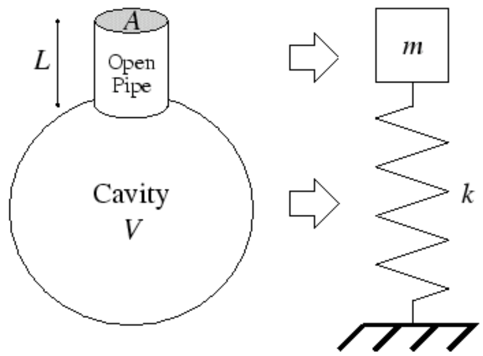


Fig. 2.2 Helmholtz resonator and its mechanical analog, a mass-spring system.

where c is the speed of sound, V is the volume of air in the container, A and L are the cross-sectional area and the effective length of the opening respectively. Therefore, the mass of the air in the neck is $\rho \times AL$, where ρ is the air density.

The effective length of the neck is greater than its geometrical length since an extra volume of air both inside and outside moves with the air in the neck. The extra length that should be added to the geometrical length of the neck is typically (and roughly) of 0.6 times the radius of the outside end, and one radius at the inside end.

In a guitar body acting as a Helmholtz resonator, the opening is the tonehole. The area of the tonehole is round and is easy to determine. The geometrical length of the neck is very short (only a couple of millimeters thick). The effective length of the neck can be approximated to about 1.7 times the radius of the tonehole. The frequency of the air resonance in a classical guitar body is often around 120 Hz [6], and is approximately a perfect fifth below that of the first plate resonance.

2.2.4 Top plate modes

Fig. 2.3 illustrates the first six modes of the top plate (predicted with Finite Element Analysis). Resonant guitar modes create large vibrations and hence radiate sound efficiently. These modes have a direct effect on the acoustic spectral response. Most guitars tend to have three body resonances in the 100-200 Hz region, due to top/back coupling and the Helmholtz mode by virtue of the soundhole. The **T(1,1)** fundamental mode (as illustrated on Fig. 2.3) usually radiates the greatest sound intensity, and the wavefronts radiate

outwards in a roughly spherical manner. The **T(2,1)** dipole radiates a volume with two large, diametrically opposing lobes. The radiation is less efficient at higher frequencies, and consequently higher frequency modes do not show as strong resonances, although they contribute to the instrument timbre.

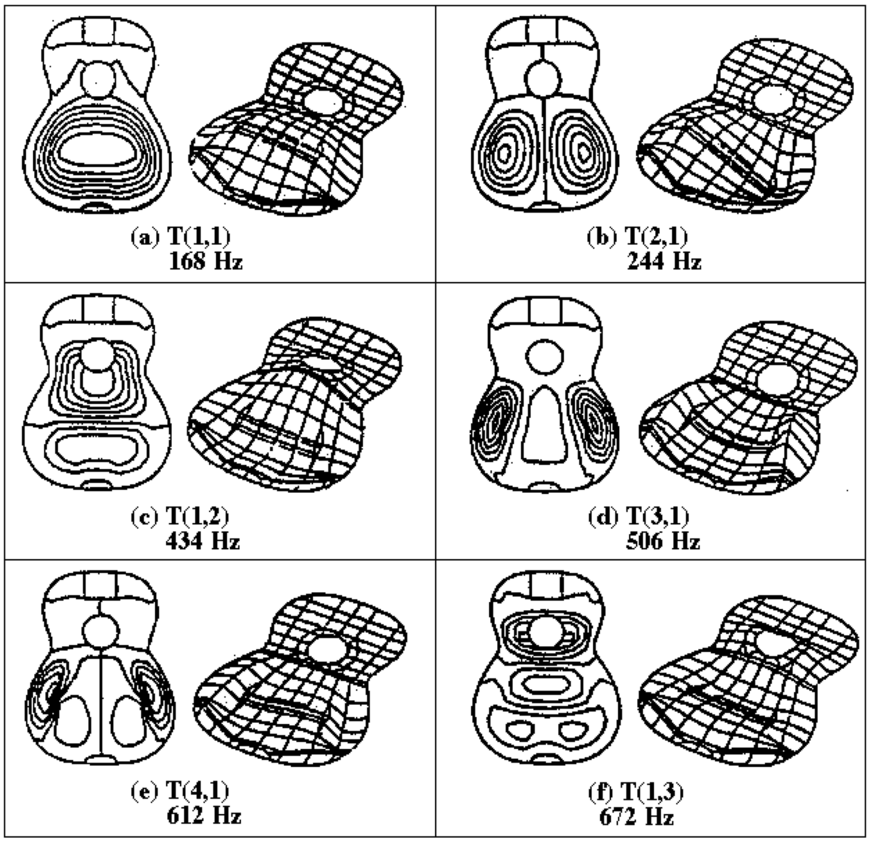


Fig. 2.3 Predicted top plate modes with the Finite Element Analysis. Figure from [2], data from [19] (in [15]).

2.3 The signal features of a guitar sound

At the moment of the attack, the player touches the string with both hands; the left principally determines pitch and the right controls loudness and timbre (for a right-handed guitarist). Timbre is, in fact, the most variable parameter within the guitarist's control.

2.3.1 The transient

The guitar body, with its own natural modes of vibration, does not immediately vibrate with the string, but responds initially in a complicated way which gives rise to the starting transient, the attack [32]. The attack is characterized by its rise time. Compared to other instruments, the guitar has an unusually quick attack.

In plucked stringed instruments, the soundboard does not start its vibrations from a state of rest; rather, it begins its motion from the shape into which it is deformed by the string, which is displaced before it is released. When the string is released from the plucker (finger or plectrum), first the top begins to vibrate in a mode that is determined by the initial deformation of the soundboard, and then it begins forced vibrations determined by the frequencies of the driving string [30].

The effect of different plucking angles on the deformation of the top plate is discussed in the next chapter.

2.3.2 The decay

The transient disappears as soon as the string has convinced the soundboard to vibrate at the string's frequency rather than of its own. In other words, a steady-state vibration is never achieved because each note begins to decay as soon as the full amplitude is reached.

Alone, the string would vibrate in a more or less regular way from the moment of release; however, its vibrations are affected by the coupling with the body through the bridge. The levels of the different partials decay at different rates, higher partials decaying faster than lower ones.

2.3.3 The spectral envelope

The main parameters affecting the spectral envelope are the choice of string, the plucking position and the direction in which the string leaves the plucking finger. Other than using a different string, the most effective method of colour modulation of a tone is to change the point at which the string is plucked [30].

The shape of the spectral envelope is at the core of this investigation on the timbre of the classical guitar and will be discussed in the next three chapters.

Chapter 3

Instrumental Gesture Parameters for the Classical Guitar

[...] Then left and right hands shall be like Male and Female Phoenix, chanting harmoniously together, and the tones shall not be stained with the slightest impurity. The movement of the fingers should be like striking bronze bells or sonorous stones. [...] These tones shall in truth freeze alike heart and bones, and it shall be as if one were going to be bodily transformed into an Immortal.
(from the description of the "Clear touch" on the Chinese lute [25]).

Contents

The sound of the guitar is determined not solely by the construction of its body, but also by the interaction between the player's finger and the string. This chapter presents the different parameters of the instrumental gesture applied by the left and right hands on a classical guitar.

We will call *instrumental gesture* the actual instrument manipulation and playing technique on an instrument [67]. We will consider here the *effective gesture* [69], defined as the purely functional level of the notion of gesture, i.e., the gesture necessary to mechanically produce the sound (like blowing in a flute, bowing on a string, pressing a key of a piano, etc.). The parameters of an instrumental gesture are, for example, the speed of an air jet, the location of a pluck along a string, or the pressure applied with a bow on a string. The variations of these parameters have an effect on the timbre and are generally clearly perceived by a trained listener such as a professional musician.

3.1 Fingering and plucking gestures

For the case of the classical guitar, there is a gesture on the left hand – the fingering gesture – and a gesture on the right hand – the plucking gesture (for a right-handed guitarist).

Fig. 3.1 Fingering and plucking gestures on the classical guitar (picture from [23] p. 9).

3.1.1 Fingering gesture

The *fingering point* on a guitar string is where a player presses a string against a fret with the tips of his left-hand fingers. The effect is a shortening of the vibrating portion of the string, determining the fundamental frequency of the tone. The fingering is therefore a *selection* as well as a *modification* gesture [68] and its parameters are the fret-string choice, the finger pressure, the vibrato amplitude and frequency, and the bending.

Fig. 3.2 Five different fingerings for an excerpt from *L'encouragement* for two guitars by Fernando Sor (1778-1839) according to guitarist Peter McCutcheon. String 1 is the highest string. Fingers are notated *p* for thumb (*pouce*), *i* for index, *m* for middle finger and *a* for ring finger (*annulaire*).

When a piece is fast and difficult, guitarists choose the most convenient fingering. Moving hands across and along the fingerboard causes qualitatively different amounts of difficulty [44]: across the neck, only the fingers are displaced and along the neck, the hand needs to be repositioned [72]. When a piece is slower, there is room for guitarists to decide on a fingering according to the timbral effects to which it leads. Fig. 3.2 shows five pos-

sible fingerings for a slow excerpt from *L'encouragement* for two guitars by Fernando Sor (1778-1839).

3.1.2 Plucking gesture

In the classical style, the string is not simply pulled aside by the fingernail. It is pushed towards the soundboard by rolling and sliding on the fingernail and is released from a position lower than its rest position having an initial amplitude and velocity distribution along its length. The string starts vibrating on a plane almost perpendicular to the soundboard so that a strong vertical force component is created at the bridge, which results in a strong soundboard response and a loud sound [15] (p. 8). The different factors that affect the string-finger interaction process are the frictional force between string and fingertip, the waves created on the string during the interaction, the physical properties of the string, and the physical properties of the finger.

While playing, every guitarist is able to vary specific parameters of the plucking action in order to obtain a desirable sound quality. These parameters are the plucking position, the pick material (finger, fingernail, plectrum), the width of the finger/fingernail/plectrum, the degree of relaxation of finger, the weight of finger on the string, and the angle with which the string is released.

The angle with which the string is released depends on the angle between finger and string (in an orthogonal plane parallel to the string) and the angle between hand and string (in an orthogonal plane perpendicular to the string).

The *plucking point* is where the player excites the string by plucking it with his or her right-hand fingers, using a pick or a fingernail. The location of the plucking point has an effect on the timbre of the tone. The plucking is therefore an *excitation* as well as a *modification* gesture. Normal plucking position is somewhere between a third and a tenth of the string length (i.e. 3 to 20 cm).

3.2 Notation for plucking techniques

The different notation systems for the plucking techniques are a unique source of information about the ways a guitarist's finger can interact with the string. The most elaborate notation system is most likely the one developed by the Chinese for the timbres of the Ch'in, an ancient seven-string lute. The notation attempts to express in words the timbre of the

tones. The terminology was borrowed from the rich vocabulary of aesthetic appreciation, used by Chinese artists and connoisseurs [25]. One of the earliest volumes, *Sixteen Rules for the Tones of the Lute*, by Leng Ch'ien (14$^{\mathrm{th}}$ century B.C.E.), describes in 150 to 200 special characters the techniques for performing the sixteen archetypical "touches" or tones of the lute. These sixteen touches are respectively described as light, loose, crisp, gliding, lofty, pure, clear, empty, profound, rare, antique, simple, balanced, harmonious, quick or slow. Rather than describe finger technique exclusively in terms of direction and strength of plucking, the Ch'in literature uses symbolic pictures to relay the "spirit" of each technique. The explanations are often accompanied by elaborate drawings. For example, the drawing of "a flying dragon grasping the clouds" (shown on Fig. 1.2) suggests that the touch should be broad and firm, the hand having more or less a clawing posture [25]. Fig. 3.3 gives an other example of a symbolic picture illustrating finger technique for playing a note on the Ch'in. All the information needed to perform a note on the Ch'in is illustrated by a single character. For example: "Kou: the middle finger pulls a string inward, 'A lonely duck looks back to the flock.' The curve of the middle finger should be modelled on that of the neck of the wild duck: curved but not angular. If the middle finger is too much hooked, the touch will be jerky." [25] (p. 127).

Several Western composers and guitarists attempted to define and notate plucking techniques more or less precisely. For example, Gilbert Biberian, for his piece *Prisms II* (1970), lists a catalogue of right-hand positions the performer should use to achieve different timbres:

- Fo. - Flautando: note is struck at the half-way nodal point;
- To. - Sul Tasto: right hand placed between 12th and 19th frets, irrespective of pitch;
- Bo. - Sul Boca: right hand placed over the sound hole;
- No. - Normale: right hand placed between sound hole and bridge, but closer to the sound hole;
- Po. - Ponticello: play as near the bridge as possible.

Another system of right-hand notation was designed by the Italian guitarist Alvaro Company in the early 1950's [22]. With his system, he aimed to expand the timbral notation of the guitar. He attempted to create a standardized right-hand notation that

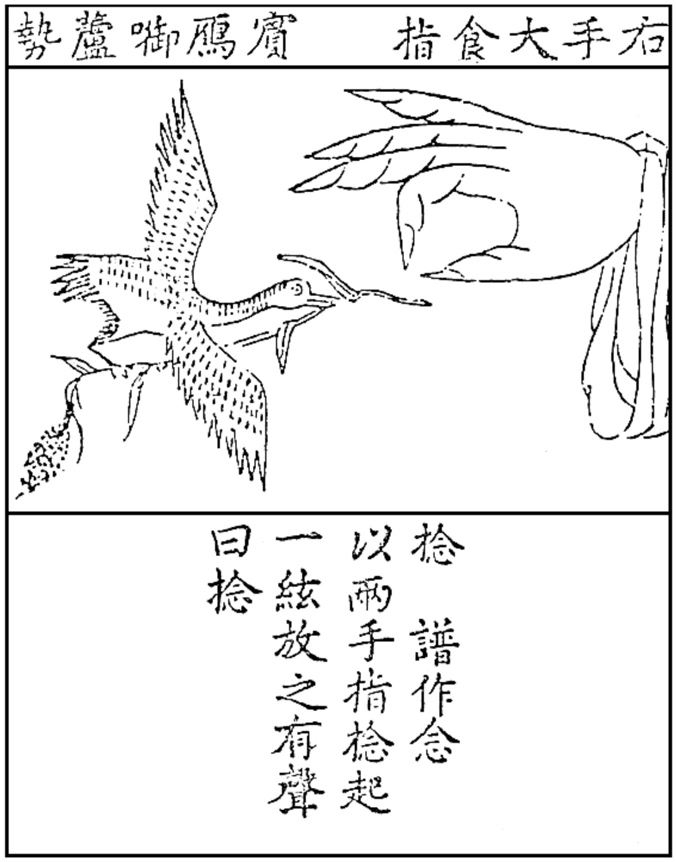


Fig. 3.3 Symbolic picture illustrating finger technique for playing a note on the Ch'in. Monumenta Nipponica Monograph, Tokyo, 1969 [25].

would take into account all aspects of right-hand technique. The great advantage of this notation system (as shown on Fig. 3.4) is that all information is transmitted in a single glance. One composite symbol indicates the player where, with what, and how to pluck the string, as do the characters in the music for the ancient Chinese Ch'in.

3.3 The main plucking parameter: the plucking position

Among the instrumental gesture parameters that contribute to the timbre of a guitar sound, the location of the plucking point along the string has a major influence. Plucking a string close to the bridge produces a tone that is softer in volume, brighter, and sharper. The sound is richer in high-frequency components. This is physically explained by considering the fact that the slope of the portion of the string connected to the bridge is steeper. On

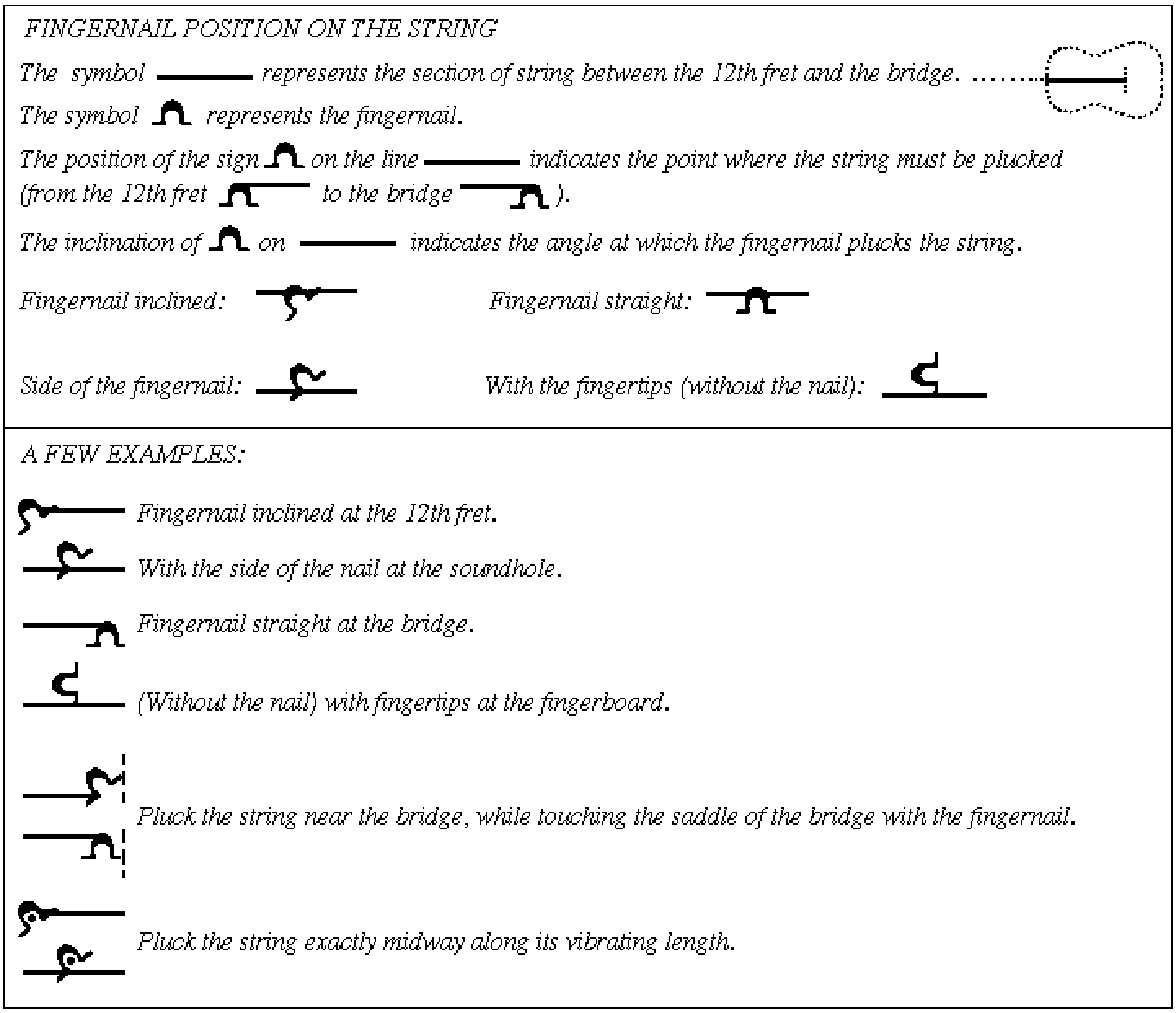


Fig. 3.4 Alvaro Company notation [22].

the other hand, plucking toward the neck (closer to the midpoint of the string) makes a louder, mellower sound, less rich in high frequency components. Because of the position of the right-hand fingers, the low strings are usually plucked further away from the bridge than the higher ones.

Sor suggests that the usual placement of the right hand should be approximately one-tenth of the whole length of the string: “For a more mellow and sustained tone, touch the string at one-eighth part of its length from the bridge... If a louder sound be desired, touch the string nearer the bridge than usual, and in this case use a little more force in touching it.” [31] (p. 4).

3.3.1 The main plucking positions

A specialized language has evolved for dealing with the description of plucking positions. This terminology is often vague since it does not refer to exact positions.

The ponticello position

Tarrega, in Gran Jota (1872), uses the ponticello position to obtain a metallic sound. In much of the early twentieth-century literature – Hindemith’s *Rondo for Three Guitars* (1925), for example – the word metallic is also used to mean ponticello [30]. Ponticello is one of the most common methods of obtaining tonal contrast in guitar music.

The tasto and flautando positions

The opposing sonority to ponticello is called sul tasto (plucking over the fingerboard) or flautando (fluted tone); these terms are also borrowed from the terminology of bowed string instruments [30]. Sor calls a “harp tone” a tone plucked halfway between the 12th fret and the bridge[1] [31], as does Tarrega/Roch: “Right hand plucks the strings at any point of the space between the 18th and the 12th frets[2], the tones are quite like those of the harp, and the more so, the higher you go” [28] (p. 69).

Musically, ponticello and tasto are often used to change the meaning of a repeating event by presenting the material in a different colour. This change of plucking position can also convey a change in the event’s character.

---

[1] The 12th fret is located at half the string’s length since an octave equals 12 semitones.

[2] The portion of the string corresponding to the 18th fret is $1/2^{18/12} = 1/2.8 \approx 1/3$.

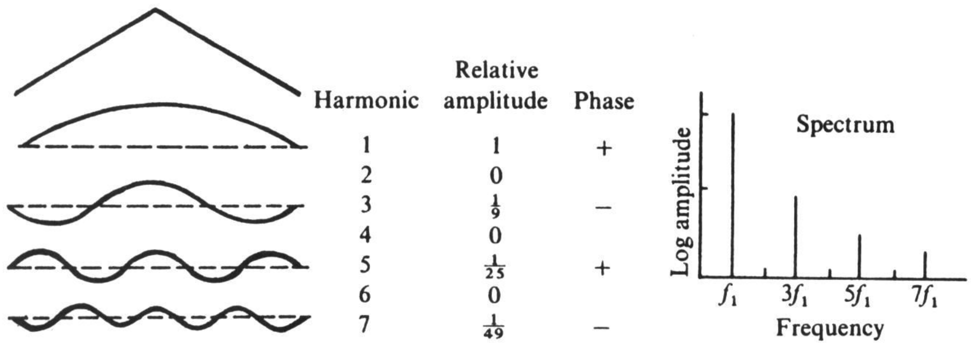


Fig. 3.5 Frequency analysis of the displacement wave of a string plucked at its midpoint. Odd-numbered modes of vibration add up in appropriate amplitude and phase to give the shape of the string [6].

The half-string tone

When a string is plucked exactly halfway along its vibrating length (above the 12th fret), a very round, harplike sound is produced. Smith-Brindle qualifies this as a "clarinet tone". The acoustical basis of this analogy is that a mid-string pluck produces only odd harmonics , which is similar to the frequency content of the tones of a clarinet, as illustrated on Fig. 3.5. The clarinet is in fact an instrument which can be approximated by a tube closed at one end and open at the other end that theoretically resonates only at odd integer multiples of the fundamental frequency.

3.4 Plucking angle and angle of release

3.4.1 Angle of release

Flamenco music needs rather short and loud tones while chamber music normally requires long duration tones. The angle of release of the string affects the coupling between the string and body modes and influences the amount of excitation of the different body modes [15]. Therefore, a player can control the balance between horizontal and vertical motion by adjusting the angle with which the string is plucked.

Classical guitarists use primarily two strokes:

- the *apoyando* stroke (also called *down*stroke or *rest* stroke);

- the *tirando* stroke (also called *up*stroke or *free* stroke).

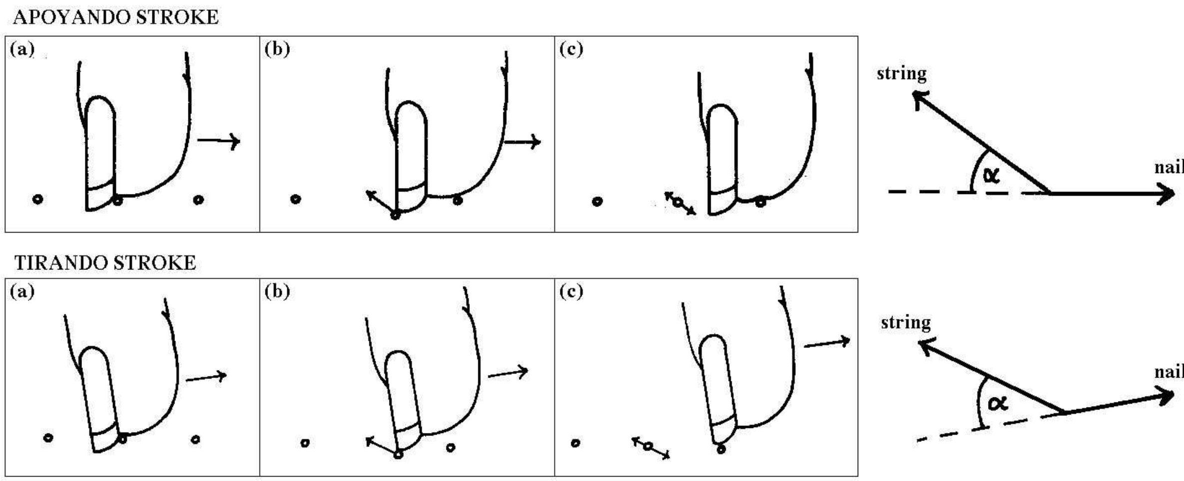


Fig. 3.6 Apoyando and tirando strokes (after [32] pp. 46-47).

Tarrega was the first teacher to develop the apoyando technique, a style of right-hand technique which calls for the fingers to be positioned perpendicular to the strings [30].

In the apoyando stroke, the finger moves parallel to the soundboard and comes to rest on an adjacent string. In the tirando stroke, the finger rises away from the strings and releases the string at a smaller angle than in the apoyando stroke.

During both apoyando and tirando strokes, the string is pushed towards the soundboard by rolling and sliding along the nail and is released from a position closer to the soundboard. The difference between the two strokes is the angle with which the string is released, as shown on Fig. 3.6. The fingernail acts as a sort of ramp, converting some of the horizontal motion of the finger into vertical motion of the string. The apoyando stroke tends to induce slightly more vertical string motion [6].

Because of its large surface and small thickness, the top plate of the guitar is more sensitive to perpendicular forces than to parallel forces. Consequently, not only do forces parallel and perpendicular to the bridge excite different linear combinations of resonances, they result in tones that have different decay rates, as shown in Fig. 3.7.

When the string vibrates in a plane almost perpendicular to the top plate, the energy is transferred to the body very efficiently and is radiated quickly into the surrounding air. The resulting tone is loud and harsh and tends to be of short duration. When the string

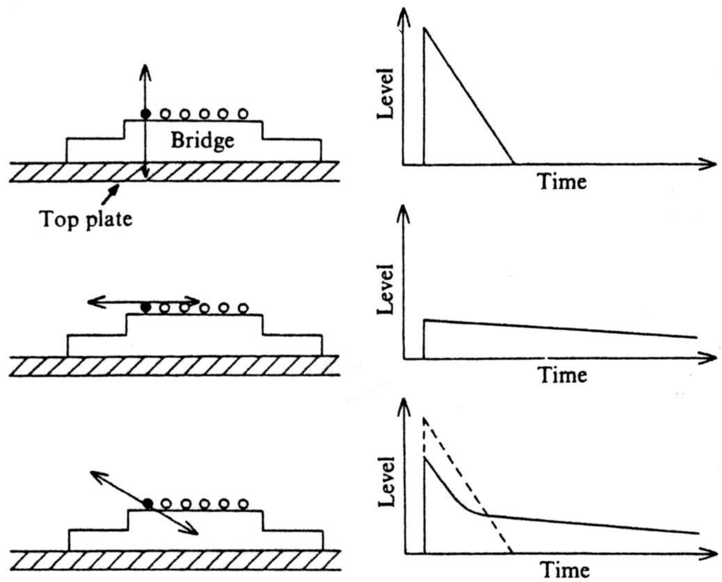


Fig. 3.7 Decay rates of a guitar tone for different plucking directions [10].

is released almost parallel to the soundboard, the sound produced is generally quieter and softer and lasts for a longer time since the energy is radiated slower. As a result, players will usually play downstroke (apoyando) for an accented tone and an upstroke (tirando) for an unaccented tone [30].

3.4.2 Effect of angle of release on the top plate modes

The angle of the fingernail's edge (ramp) is very important in determining the speed and direction with which the string will travel as it leaves the finger. Jansson defines a three-coordinate system centred on the bridge in order to decribe the plucking direction. As shown on Fig. 3.8:

- the x axis is the axis parallel to the soundboard and perpendicular to the strings;
- the y axis is the axis parallel to the soundboard and parallel to the strings;
- the z axis is perpendicular to the soundboard and perpendicular to the strings.

Both the angle of the finger or nail in the $x-y$ plane and the angle of the string displacement in the $x-z$ plane alter the spectrum of a tone [9]. Jansson has shown the

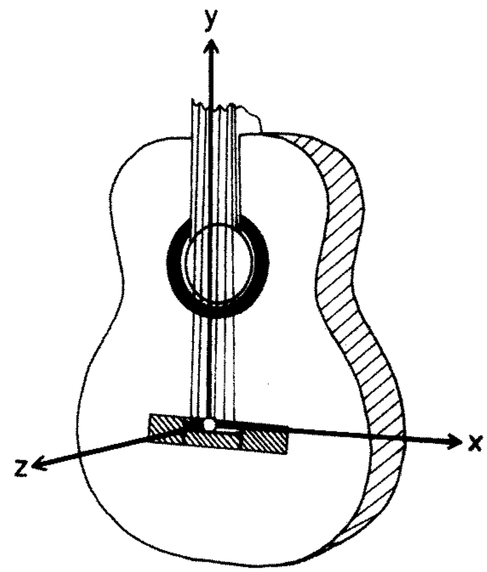


Fig. 3.8 Coordinate system for the guitar angle [9].

Top Displacement (TD) modes for forces applied in these three directions.

- TD1 occurs when the string is displaced in the z-direction: the bridge vibrates as a whole piece along this direction. TD1 corresponds to T(1,1) on Fig. 2.3. Typical values would be around 150 Hz.
- TD2 occurs when the string is displaced in the x-direction: the bridge pivots about its middle and around an axis parallel the the string, its two edges being in alternate positions like a swing. TD2 corresponds to T(2,1) on Fig. 2.3. Typical values would be around 235 Hz [30].
- TD3 occurs when the string is displaced in the y-direction: the bridge pivots around an axis parallel to its length (perpendicular to the string). This top displacement is negligible in comparison to the other two modes because one needs at least four times the force that it takes to produce TD1. TD3 corresponds to T(1,2) on Fig. 2.3.

Moving the string in the z-direction creates combinations of TD1 and TD2, especially if the string that is displaced is far away from the middle of the bridge. Most displacements can be described by the string's movement in a combination of the x, y and z directions,

so that the resulting top deformations will be combinations of the three modes TD1, TD2, and TD3 [9]. On Fig. 3.9, the solid black line depicts the actual shape of the soundboard.

If the displacement is in the x-direction, the TD2 appears (case (I) on Fig. 3.9). Plucking one of the lower strings tirando (upstroke) with the thumb at approximately 30 degrees (case (II) on Fig. 3.9) gives the combination of TD1 and TD2, where the total deformation can be decomposed into

(a) TD2, depending on the x component,

(b) TD2, depending on the z component,

(c) TD1, depending on the z component.

The next example (case (III) on Fig. 3.9) illustrates what happens when the same string is plucked apoyando (downstroke). The z-component is then negative, which changes (b) to a negative value, and the resulting combination of **(a)** and **(b)** is a much smaller value for TD2, so the prefix of that note will contain much less of that mode [30].

3.4.3 Effect of angle on attack

The frequencies of TD1 and TD2 along with the Helmholtz mode (air mode A_0) are present in the attack of a guitar note. Those frequencies are generally not harmonically related to the fundamental of a tone. The air mode is usually between two fretted notes on the guitar; whenever either of these notes is played, the vibrations of the top plate excite this mode and that frequency is strongly reinforced from within the instrument. Moreover, each time TD1 is excited, the air mode becomes a part of the sound produced.

The amount of noise in the transient of a note varies with the angle of the string's displacement before its release. It is also the case that the further away from the bridge the string is plucked, the less energy is put into these noise elements. A change in the angle of string displacement also changes the amount of air resonance in the transient. This is because the TD1 and the air resonance are so closely linked. If the pluck is perpendicular to the soundboard, the air mode is much more present. This effect occurs regardless of which string is plucked and how it is fretted. According to Schneider [30], the fact that the amout of air and the TD modes stay the same for a given plucking angle is one of the factors that provides timbral continuity, telling the ear that the same "instrument" is playing when a melody or scale crosses strings or octaves.

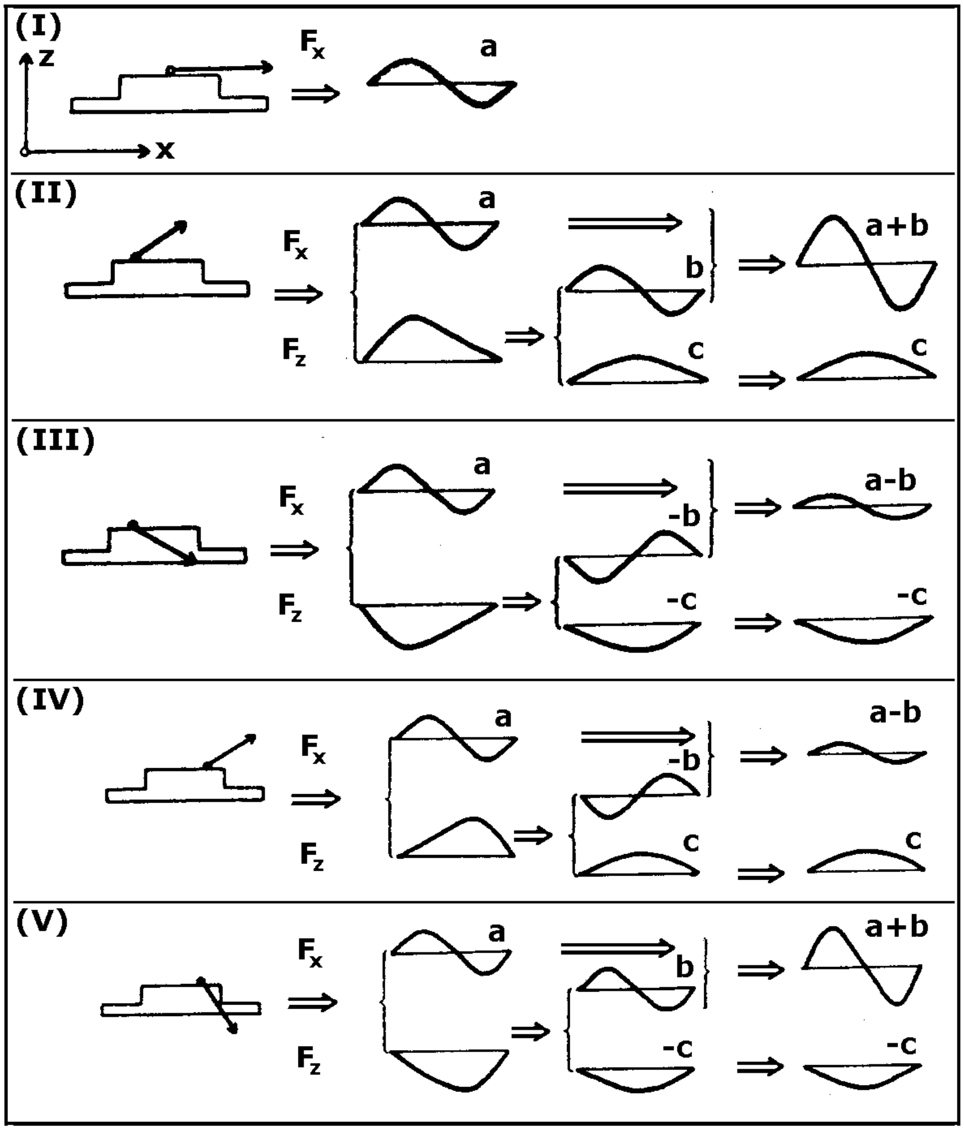


Fig. 3.9 Qualitative behaviour of the soundboard when a force is applied transversely **(I)** to one of the higher four strings at 0^o; **(II)** to one of the lower three strings at 30^o; **(III)** to one of the lower strings at -30^o; **(IV)** to one of the higher three strings at 30^o; **(V)** to one of the higher three strings at -30^o. The top displacements are illustrated with the profiles notated **(a)** for TD2 depending on the x component, **(b)** for TD2, depending on the z component, **(c)** for TD1, depending on the z component [9].

3.5 Effect of plectrum width

3.5.1 Lowpass filtering due to plectrum width

The plectrum acts as a low-pass filter: the thinner the width, the higher the cutoff frequency [30]. In fact, modes of vibration with a wavelength shorter than twice the plectrum width are very slightly excited and their frequencies are almost absent from the sound spectrum[3]. In other words, widening the plectrum, whether with flesh, nail or plastic, has the effect of damping the higher harmonics, thus producing a less bright, sweeter sound. This occurs because the edges of the force waveform are rounded by the change in the initial curve of the displaced string [30].

3.5.2 Changing plectrum width by changing angle

A popular method among performers of changing the width of the plectrum consists of altering the angle with which the finger approaches the string (i.e. the angle of attack) which is defined as the angle between the line of the hand's knuckles and the string length [32].

For instance, when the line of knuckles is set parallel to the strings, the angle of attack is equal to 0 degrees. Guitarists claim that when the nail is turned at a larger angle in relation to the string, the sound changes from thin to warm. Consequently, by altering the angle of attack, the performer uses plectra of different widths since the string comes in contact with a larger or smaller area of the fingernail, depending on the angle.

The lowpass filtering accompanying an increase in the plectrum width by changing angle is illustrated on Fig. 3.10 and 3.11.

3.6 Plucking with finger, nail or pick

Pavlidou created a three-dimensional physical model of the string-finger interaction [15]. The simulations predict the movement of the string and fingertip during the interaction, the amplitude and velocity distributions of the string upon release, the force waveform on the bridge and the subsequent free string vibrations [15]. Results from the computational model show that the string-finger interaction is strongly influenced by the frictional characteristics

[3] For example, assuming the sound of a transverse wave travelling at 50 m/s along the string, if the plectrum width w is 2 mm, the shortest wavelength is 4 mm and the cutoff frequency is $f_{max} = c/\lambda = c/2w = 50/0.004 = 12500$ Hz.

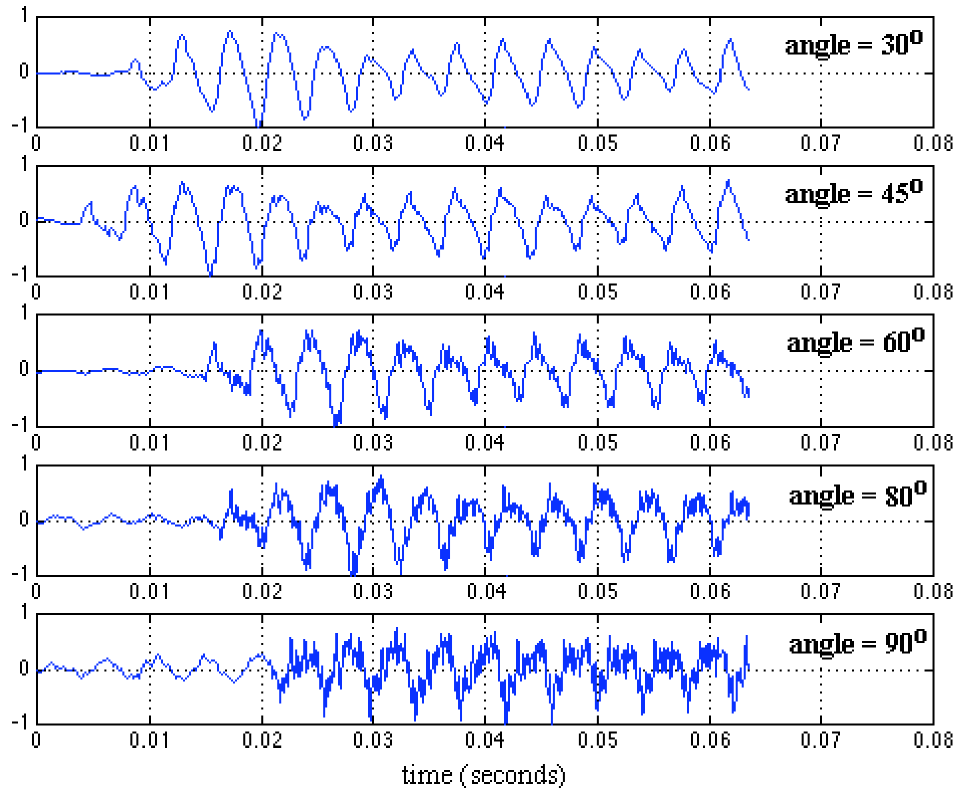


Fig. 3.10 First 70 ms of the acoustic signal of B-string plucked 18 cm away from the bridge with different angles. 90^o corresponds to the plucking finger perpendicular to the string.

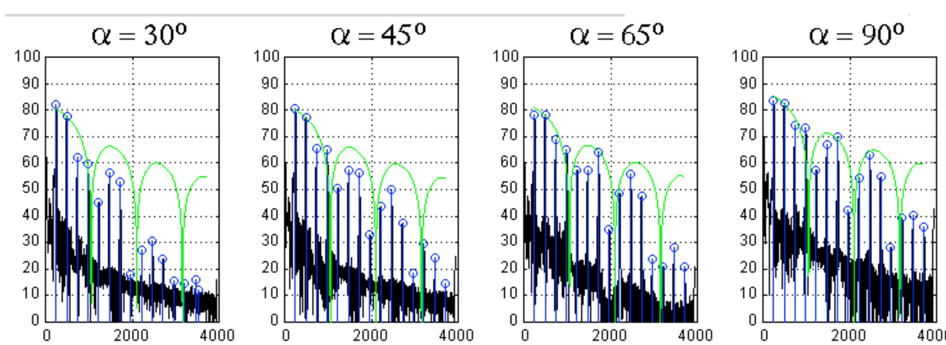


Fig. 3.11 Magnitude spectrum (dB vs Hz) of B-string plucked 18 cm away from the bridge with different angles. 90^o corresponds to the plucking finger perpendicular to the string. Theoretical spectral envelope is superimposed on the magnitude spectrum. Spectral tilt is correlated with plucking angle.

of the fingernail, the response of the finger-muscle, the input admittance of the body and the direction of the finger movement.

The choice of plectrum affects the sound because its thickness determines the cut-off frequency of the string vibrational modes. Santisteban describes: "To obtain a full and mellow tone, apply some force with the ends of the fingers. As the finger leaves the string, the nail will come into contact witht the string producing a rich tone. In order to produce a brittle sound, use only the nail in producing the sound" [29].

When plucking the string with the flesh of the fingertip, which corresponds to a thick and soft plectrum, the sound is full and its spectrum contains only low-frequency harmonics. When only the nail is used for plucking the string, the sound is thinner and its spectrum contains high frequency harmonics.

The classical style of guitar playing requires that the nail, rather than the flesh of the fingertip, be used to pluck the string. It was Tarrega who first introduced the use of the nail in the guitar.

3.6.1 Playing with nail

The use of nail brightens the guitar tone since it acts as a sharp plectrum and excites also the high frequency vibrational modes of the string. Guitar performers state that by using the nail, they have better control over the string during the interaction, because they can more readily predict the moment at which the string will be released.

The shape of the fingernail is always finely adjusted by filing and not cutting, while the length of the nail is adjusted in such a way that when the hand takes its position in relation to the guitar, each nail is placed at the same distance from the string.

3.6.2 Frictional characteristics of the nail and travelling waves

The string-finger interaction is a dynamic process which involves friction between the fingertip of the player and the string. In the beginning of the interaction the string sticks or rolls along the nail due to the friction between them. The string starts slipping along the nail when the friction reaches its maximum value [15].

The string's trajectory during the interaction, the exact point at which the string leaves the finger, the velocity of the string on release and the duration of the interaction are all determined by the physical parameters of the string (such as tension, density, stiffness,

shear modulus, etc.), the fingertip (such as nail shape, mass, frictional characteristics, etc.) and the local forces exerted on the string during the interaction process.

Moreover, during the interaction time, longitudinal, transverse, and torsional waves are created on the string and travel along its length. After their reflection by the two ends of the string and upon their return to the plucking point, they find the string still in contact with the fingertip; their existence alters the local conditions and determines the future movement of the string and the fingertip.

It must be noted that the waves reflected by the bridge end of the string are not the exact reverse of the incoming waves since the bridge's own movement modifies them. The other end (i.e. the nut of the string) also modifies the incoming waves, but to a lesser extent since it is almost perfectly rigid.

When the modified reflected waves return to the plucking position carrying information from the guitar body, the fingertip, still in contact with the string, is able to detect and evaluate this information. Experienced players, when selecting an instrument to purchase, touch and interact with the guitar string without releasing it in order to evaluate the information from the body [15].

3.6.3 Stick-slip motion of the string during the string-finger interaction

Similarities can be found between the string-finger interaction for the guitar and the string-bow interaction for the violin [15]. With the guitar, the frictional forces occuring during the interaction between string and fingertip are similar to those of the interaction between the bow and the string with the violin, producing a stick-slip motion of the string. The string element which touches the fingertip rolls and sticks on the fingernail until the friction between them reaches a critical value. After this point, the string element starts slipping along the fingernail and finally leaves it to vibrate freely. The difference with the violin is that the stick-slip motion only occurs during a very short amount of time before the string is released [15].

3.7 Articulation

The term articulation refers, in music, to the manner in which tones are attacked and released. According to Duncan, the mastery of articulation goes to the heart of mastering

an instrument's way of producing sound [23]. The term phrasing pertains more to the manner in which tones are grouped for expressive purposes.

Guitar tones can have various articulations: *martelé*, *spiccato*, *détaché*, or *staccato*. The articulation has to do with a performer's control of note length, irrespective of written rests. Playing *staccato* reduces nominal note value by more than half; it is the shortest note. Playing *legato* gives notes their full value and joins the notes without a perceptible break. A true legato is impossible on the guitar. The nature of the instrument – that necessarily entails a percussive mode of attack followed by a rapid note decay – produces consecutive articulations. Duncan depicts the difference between *legato* and *staccato* with the difference between the word *oar* and the word *toe* when repeated in sequence.

Articulation also refers to the degree of percussiveness in the attack, particularly with the technique of wind and string instruments. "It is [also] more akin to the effect that different consonants have upon the same vowel sound in speech" [23]. On the violin, *martelé* is a percussive stroke with a consonant type of sharp accent at the beginning of each stroke and always a rest between strokes.

Duncan adds that "articulation pauses before notes allow control of color and of rhythmic placement. They enhance the clarity of one's musical enunciation by providing space for notes to breathe". As Duncan explains, guitarists often use words related to speech to describe their playing techniques : "consonant type of sharp accent", articulation, clarity, enunciation, breath, etc.

3.8 Vibrato

A guitar note inevitably changes throughout its duration, not only in loudness but in quality, since the partials decay at different rates. Though the guitar is far from unique in producing notes which decay gradually and change in the process, the possibility of vibrato distinguishes it from instruments such as the piano, the harpsichord and the harp.

Vibrato is a periodic variation of the fundamental frequency of the note. It is usually accompanied by synchronous pulsations of loudness and timbre [97]. On the violin, vibrato is accomplished by altering the length of the string. On the guitar, however, because it is a fretted instrument, this frequency modulation must be achieved by altering the string tension and hence the pitch. For notes above the 5^{th} fret, the technique usually consists of pushing and pulling the string toward and away from the bridge; for those notes that

lie closer to the nut, the string is pulled from side-to-side, perpendicular to the other strings [30].

3.8.1 Vibrato rate

Orchestral players have been found to favour a vibrato rate of 6 or 7 Hz. This is also the natural rate at which singers modulate the voice [146]. The vocal vibrato develops more or less automatically during voice training [139] and is the result of the intermittent supply of nerve energy to the mechanism (at the frequency of stammering and other spasmodic movements) [147]. It may be that the use of vibrato appeals by imbuing the instrumental sound with a vocal quality.

3.8.2 Vibrato frequency range

The range of the pitch variation is usually about a quarter-tone either side of the note with singers (Seashore [97] measured an average extent of ±48 cents), but only half that amount with violonists. This width of vibrato is mostly a matter of taste and fashion.

3.8.3 Perceptual effect and musical function of vibrato

A vibrato of about the optimum frequency and of moderate width is not experienced as a variation in pitch, but is rather perceived as a rich and warm quality, bringing life to the tone [32]. Seashore states that it gives a pleasing flexibility, tenderness, and richness to the tone [97]. Musically, vibrato is used to accentuate phrase endings, to make individual melodic notes stand out from their neighbours or to highlight the emotional content of the piece. This technique of tone modification was thoroughly described by the Chinese lute masters, who called it the "Loose Touch" and who ranked it among the sixteen important aspects of tone production [25].

The vibrato allows the "sweeping" of the spectral envelope, thereby adding to the vocal quality of guitar sounds. The Russian historian Makaroff described a Spanish guitarist with very evocative terms : "The vibrato, when performed by Ciebra, was really divine – his guitar actually sobbed, wailed and sighed. Ciebra only showed these remarkable qualities in slow tempos as in largo, adagio or andante." [21].

The spoken voice seldom has vibrato; it is nonetheless always inflected: a definite pitch is almost never sustained. In fact, some vowels are more recognizable when inflected than

when not [138]. Inflection and vibrato are both variations of the fundamental frequency, inducing a “sweeping” of the spectral envelope which eases the recognition of the sound.

Chapter 4

The Physics of the Plucked String

Contents

This chapter describes the physical behaviour of the plucked string. The magnitude spectrum coefficients of an ideal plucked string are derived (for the displacement, velocity and acceleration waves). Finally, differences between an ideal string and a real string are presented.

4.1 Standing waves on an ideal string

When a string is plucked, two pulses or waves are sent travelling in opposite directions down the length of the string (Fig. 4.1 on the left). When each of these travelling waves reaches the string's boundary, it is reflected back again in the opposite direction, inverted (Fig. 4.1 on the right). The waves travelling on the string are mostly transverse, but there are also longitudinal and torsional waves.

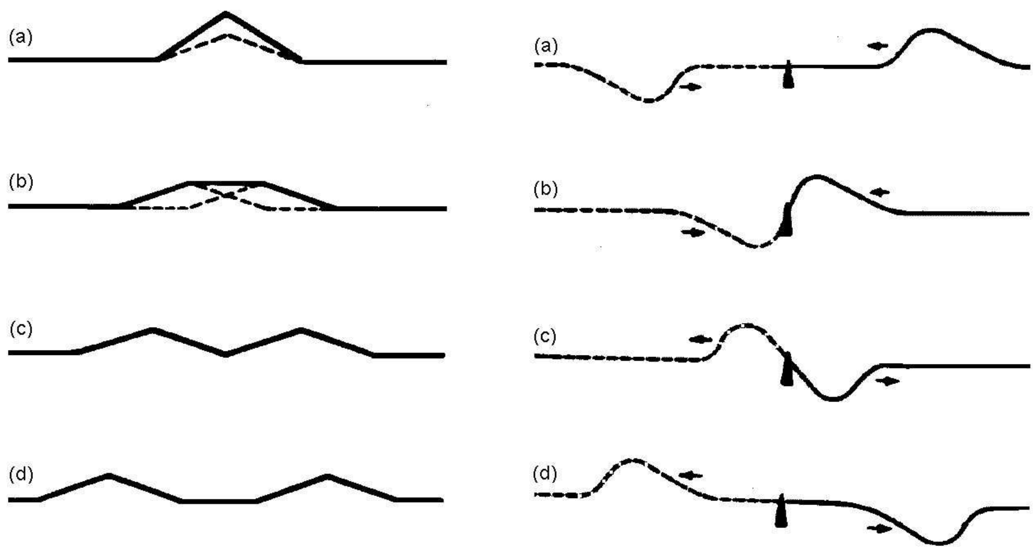


Fig. 4.1 On the left : motions of a plucked string. The solid lines give the shapes of the strings at successive times, and the dotted lines give the shapes of the two (backward and forward) travelling waves, whose sum is the actual shape of the string. On the right : reflection of a wave from the end support of a string. In this case, the dotted lines show the imaginary extension of the waveform beyond the end of the string [14] (pp. 75-76).

An excitation, such as a pluck, in a real physical string initiates wave components that travel independently in opposite directions (dashed curves on Fig. 4.2). The resulting motion consists of two bends, one moving clockwise and the other counterclockwise around a parallelogram. The output from the string, that is the force at the bridge of an acoustic instrument or the pickup voltage in an electric guitar, reacts to both wave components.

Each wave then travels to the other end of the string, where the process is repeated. Since these two travelling waves are moving on the same string, they cross and interfere

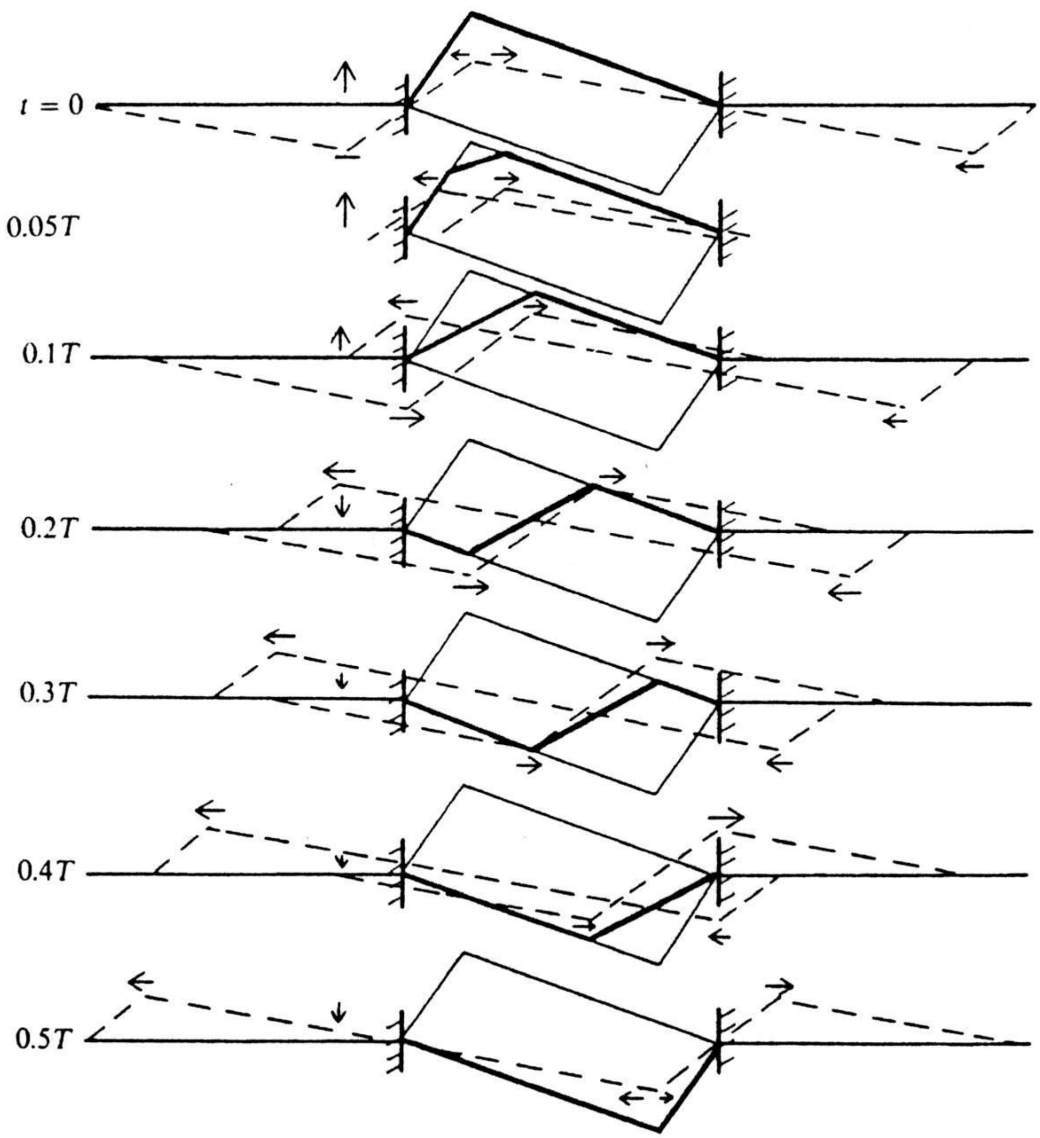


Fig. 4.2 Time analysis through one half cycle of the motion of a string plucked 1/5th of the distance from one end [6].

with each other as they travel from one end of the string to the other. Their amplitudes are added together at all points; if, at a certain point, both waves are positive, the combined value will be larger than that of either one alone. If, at another point, one is positive and the other negative, they will cancel each other out so that the combined value is zero. The result of this superposition of waves is a standing wave.

4.2 Missing harmonics in a plucked string spectrum

As illustrated on Fig. 4.3, when a string is set into vibration with a pluck, the sound signal lacks the harmonics that have a node at the plucking point. For example, plucking a string at its middle (L/2) prevents the even partials from being initiated. On the other hand, a partial is initiated maximally at this antinodal position(s).

4.3 Time and frequency analysis of plucked string

In the string model considered in this section, the string is assumed to be ideal (i.e. with no stiffness and no damping), displaced from its rest position to an inital shape, and released with zero initial velocity along its length.

This simple description of the plucked string explains to some extent how different performers produce a variety of sounds in a guitar, namely by altering the plucking position along the string. However, the idealized plucked string description cannot explain how a guitarist, while using a steady plucking position and plectrum, is able to produce a variety of different sounds on the same guitar.

4.3.1 The transverse wave equation

It is assumed that only transverse waves travel along the string. Let $y(x,t)$ the vertical displacement of an ideal string of length l with fixed ends as a function of the position along the string x ($x = 0$ is at the bridge termination and $x = l$ is at the nut termination for example) and as a function of time t. The string motion takes place only on the xy-plane and it can be described through the one-dimensional version of the wave equation which was first derived in 1747 by D'Alembert for the case of the vibrating string [3]. This

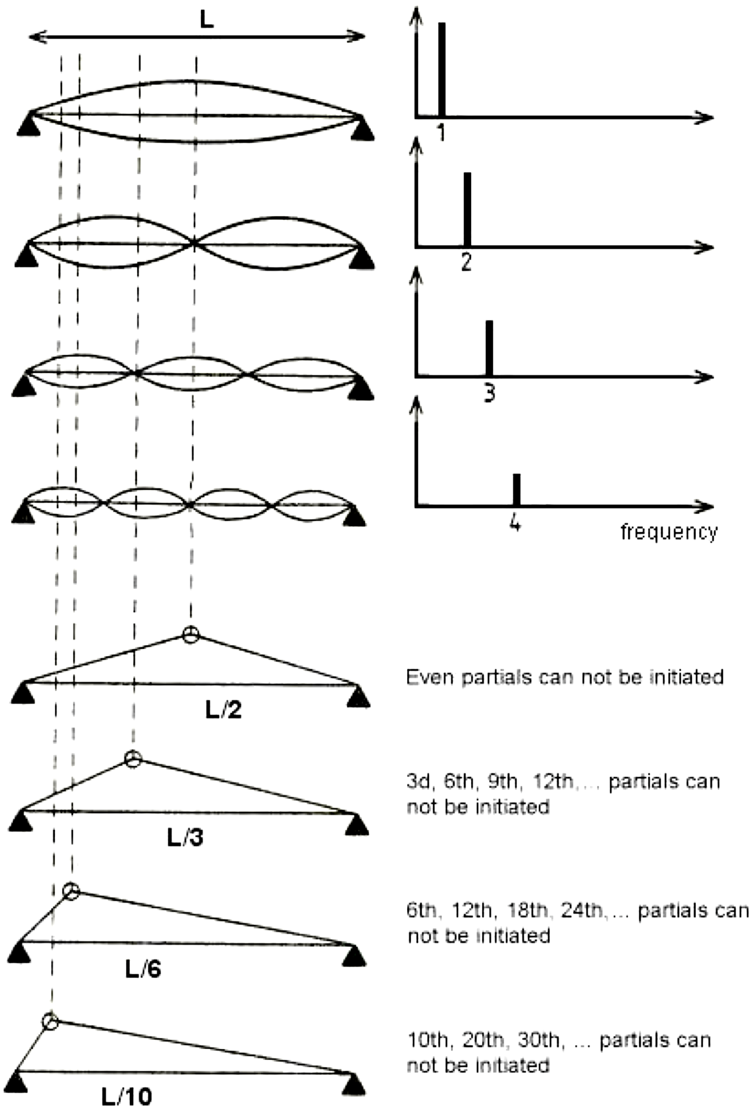


Fig. 4.3 Demonstrations of the influence of plucking position. A partial can not be initiated at the nodal position of the corresponding standing wave [10] (p. 14).

equation is known as the transverse wave equation :

$$\frac{\partial^2 y(x,t)}{\partial t^2} = c^2 \frac{\partial^2 y(x,t)}{\partial x^2} \tag{4.1}$$

where

$$c = \sqrt{T/\mu} \tag{4.2}$$

is the speed of propagation of the transverse wave on the string, square root of the ratio of T, the tension (in N or kgm/s^2) and of μ, the mass per unit length of the string material (in kg/m).

The two string ends are assumed to be fixed during the vibration of the string, as described by the conditions

$$y(0,t) = y(l,t) = 0 \tag{4.3}$$

For the ideal string of length l with rigid end supports, the frequencies f_n of its vibrational modes are multiple integers of the fundamental frequency f_o, given by

$$f_n = nf_0 = n\frac{c}{2l} = \frac{n}{2l}\sqrt{\frac{T}{\mu}} \tag{4.4}$$

where n is the order of the partial. The frequency increases if the tension increases, or if the length is shortened, or if the mass per unit length decreases. In the ideal case, there is an infinite number of normal modes, which results in an infinite series of harmonics in the spectrum of the sound.

The most general integral solution of Eq. (4.1) which fulfills the conditions of Eq. (4.3) and corresponds to a periodic motion of the string can be written as the sum of normal modes [6]:

$$y(x,t) = \sum_{n=1}^{\infty} (A_n \cos \omega_n t + B_n \sin \omega_n t) \sin\left(\frac{n\pi x}{l}\right) \tag{4.5}$$

where A_n and B_n are constant coefficients which can be determined from the shape and velocity of the string for any given time t and $\omega_n = n\omega_0 = n(2\pi f_0)$. At time $t = 0$, the shape of the string is given by

$$y(x,0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{l}\right) \tag{4.6}$$

and the velocity by

$$v(x,0) = \left.\frac{dy(x,t)}{dt}\right|_{x=0} = \sum_{n=1}^{\infty} n\omega_0 B_n \sin\left(\frac{n\pi x}{l}\right) \tag{4.7}$$

4.3.2 Initial displacement conditions

An ideal plucking excitation is a static displacement and then an abrupt release of the string at one particular point. The string is initially pulled aside at $x = p$ by a sharp point in such a way that, at $t = 0$ when it is released, it forms two straight lines proceeding from the plucking position to the fixed ends.

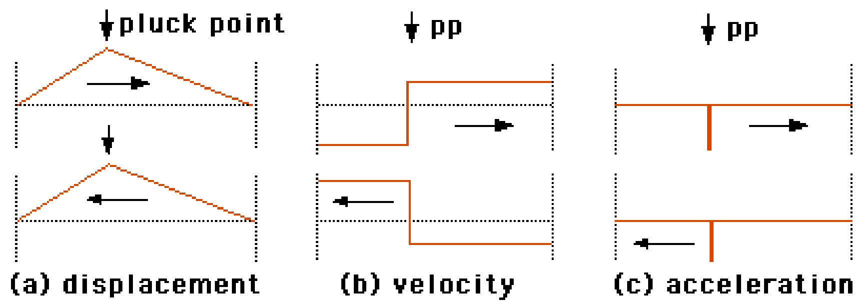


Fig. 4.4 Plucked string behaviour immediately after an ideal pluck for **(a)** displacement, **(b)** velocity and **(c)** acceleration waves [46].

Fig. 4.4 illustrates the initial conditions of a string after the release. For each wave variable, the backward and forward travelling waves are represented. To obtain the actual initial conditions, the two waveforms are added. The displacement waveforms **(a)** are triangular, the velocity waves **(b)** are step functions (their sum is null), and the acceleration waves **(c)** are impulse-like.

An ideal plucking excitation at a distance p from an end and with amplitude h is such that all points along the string have a zero initial velocity:

$$v(x,0) = \dot{y}(x,0) = 0 \qquad \text{for all x,} \tag{4.8}$$

and the string is initially shaped like a triangle with its summit at point (p, h):

$$y(x,0) \quad = \quad \frac{h}{p}x \qquad \text{for } 0 \leq x \leq p \tag{4.9}$$

$$y(x,0) = \frac{h(l-x)}{l-p} \quad \text{for } p \leq x \leq l \tag{4.10}$$

With these initial conditions, the coefficients A_n and B_n can be calculated from their expression:

$$A_n = \frac{2}{l}\int_0^l y(x,0)\sin\left(\frac{n\pi x}{l}\right)dx \tag{4.11}$$

$$B_n = \frac{2}{\omega_n l}\int_0^l \dot{y}(x,0)\sin\left(\frac{n\pi x}{l}\right)dx \tag{4.12}$$

Because of the zero initial velocity (Eq. 4.8),

$$B_n = 0.$$

Hence, the amplitude of the nth mode of the vertical displacement wave y is

$$C_y[n] = \sqrt{A_n^2 + B_n^2} = |A_n| \tag{4.13}$$

where A_n is obtained by solving by parts the integral in Eq. (4.11) :

$$\begin{aligned}
A_n &= \frac{2}{l}\left(\int_0^p \frac{h}{p}x\sin(n\pi x/l)dx + \int_p^l \frac{h(l-x)}{l-p}\sin(n\pi x/l)dx\right) \\
&= \frac{2}{l}\frac{h}{p}\left[-\frac{x}{n\pi/l}\cos(n\pi x/l) + \frac{1}{(n\pi/l)^2}\sin(n\pi x/l)\right]_0^p \\
&\quad + \frac{2}{l}\frac{hl}{l-p}\left[\frac{-\cos(n\pi x/l)}{n\pi x/l}\right]_p^l - \frac{2}{l}\frac{h}{l-p}\left[\frac{-\cos(n\pi x/l)}{n\pi x/l} + \frac{\sin(n\pi x/l)}{(n\pi/l)^2}\right]_p^l \\
&= -\frac{2h}{n\pi}\cos(n\pi p/l) + \frac{2h}{lp(n\pi/l)^2}\sin(n\pi p/l) + \frac{2h}{l-p(n\pi/l)}\cos(n\pi p/l) \\
&\quad - \frac{2h}{l(l-p)}\frac{p}{(n\pi/l)}\cos(n\pi p/l) + \frac{2h}{l(l-p)(n\pi/l)^2\sin(n\pi p/l)} \\
&= \frac{2h}{l(n\pi/l)^2}\left(\frac{1}{p} + \frac{1}{l-p}\right)\sin(n\pi p/l) \\
&\quad + \left(-\frac{1}{n\pi} + \left(\frac{1}{(l-p)(n\pi/l)} - \frac{p}{l(l-p)(n\pi/l)}\right)\right)\cos(n\pi p/l)
\end{aligned}$$

Finally,

$$A_n = \frac{2h}{n^2\pi^2 p(l-p)/l^2}\sin(n\pi p/l) \tag{4.14}$$

and the amplitude of the nth mode of the vertical displacement wave y is expressed by

$$C_y[n] = \frac{2h}{n^2\pi^2 R(1-R)}|\sin(n\pi R)| \tag{4.15}$$

where

$$R = p/l \tag{4.16}$$

is the relative plucking position, defined as the fraction of the string length from the point where the string was plucked to the bridge.

The equation giving the vertical displacement of the string as a function of the position x, of time t and of the plucking relative position R (Eq. 4.5) becomes

$$y(x,t) = \sum_n \left(\frac{2h}{n^2\pi^2 R(1-R)}\sin(n\pi R)\right)\cos(\omega_n t)\sin\left(\frac{n\pi x}{l}\right) \tag{4.17}$$

4.3.3 Displacement, velocity, acceleration and force waves

Knowing the string movement, the vertical force $F(t)$ exerted on the bridge by the string can be calculated from the string slope near the bridge, as

$$F(t) = T\frac{\partial y(x,t)}{\partial x} \tag{4.18}$$

The force waveform is a pulse with duty cycle $(1/R-1)$. In fact, the ratio of the durations of the positive and negative segments of the force waveform is equal to $(1/R-1)$. For example, if $R = 1/5$, duty cycle ratio $= 4$. This is the case (b) on Fig. 4.5. Now, in order to obtain the equation for the velocity variable, the derivative of Eq. 4.17 is taken with respect to time.

$$v(x,t) = \frac{\partial y(x,t)}{\partial t}$$

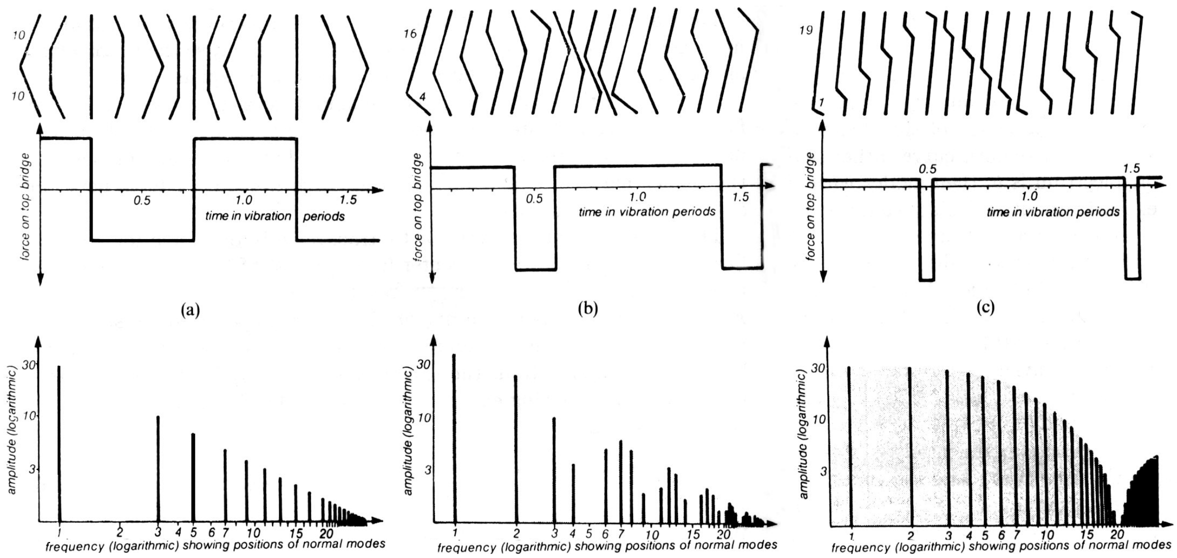


Fig. 4.5 On the top of the figure are shown the string shapes at successive intervals during the vibration period, for a string plucked at its center (a), at 1/5 of its length (b), at 1/20 of its length (c) from the bridge. In the middle, pulse-shape waveforms of transverse bridge force are displayed. At the bottom of the figure are the corresponding spectra [5].

$$
\begin{aligned}
v(x,t) &= \sum_n \left(\frac{2h}{n^2\pi^2 R(1-R)} \sin(n\pi R) \right) (-\omega_n \sin(\omega_n t)) \sin\left(\frac{n\pi x}{l}\right) \\
&= \sum_n \left(\frac{(2h)(-2\pi n f_0)}{n^2\pi^2 R(1-R)} \sin(n\pi R) \right) \sin(\omega_n t) \sin\left(\frac{n\pi x}{l}\right) \\
&= \sum_n \left(\frac{-4hf_0}{n\pi R(1-R)} \sin(n\pi R) \right) \sin(\omega_n t) \sin\left(\frac{n\pi x}{l}\right)
\end{aligned}
$$

And similarly for the acceleration variable:

$$
a(x,t) = \frac{\partial v(x,t)}{\partial t}
$$

$$
\begin{aligned}
a(x,t) &= \sum_n \left(\frac{-4hf_0}{n\pi R(1-R)} \sin(n\pi R) \right) \omega_n \cos(\omega_n t) \sin\left(\frac{n\pi x}{l}\right) \\
&= \sum_n \left(\frac{(-4hf_0)(2\pi n f_0)}{n\pi R(1-R)} \sin(n\pi R) \right) \cos(\omega_n t) \sin\left(\frac{n\pi x}{l}\right) \\
&= \sum_n \left(\frac{-8hf_0^2}{R(1-R)} \sin(n\pi R) \right) \cos(\omega_n t) \sin\left(\frac{n\pi x}{l}\right)
\end{aligned}
$$

Let

$$K(R) = \frac{2h}{R(1-R)} \tag{4.19}$$

$K(R)$ is a constant for a given R. The magnitude of the spectral components becomes:

- for the displacement variable:

$$C_y[n] = \frac{K(R)}{n^2\pi^2} |\sin(n\pi R)| \tag{4.20}$$

- for the velocity variable:

$$C_v[n] = \frac{2K(R)f_o}{n\pi} |\sin(n\pi R)| \tag{4.21}$$

- for the acceleration variable:

$$C_a[n] = 4K(R)f_o^2 |\sin(n\pi R)| \tag{4.22}$$

The sine term at $(n\pi R)$ in Eq. (4.20), (4.21), (4.22) allows no energy at the $1/R$th harmonic frequency nor at integer multiples of that frequency (since $sin(n\pi R)$ equals 0 when the product nR is an integer).

The expressions for $C_y[n]$, $C_v[n]$ and $C_a[n]$ can be interpreted as spectral envelopes if the discrete integer variable n (the order of the partial) is replaced by a continuous variable f/f_o where f_o is the fundamental frequency. For example, for the displacement wave, the expression becomes:

$$C_y(f) = \frac{f_o^2 K(R)}{f^2\pi^2} \left| \sin\left(\frac{\pi f R}{f_0}\right) \right| \tag{4.23}$$

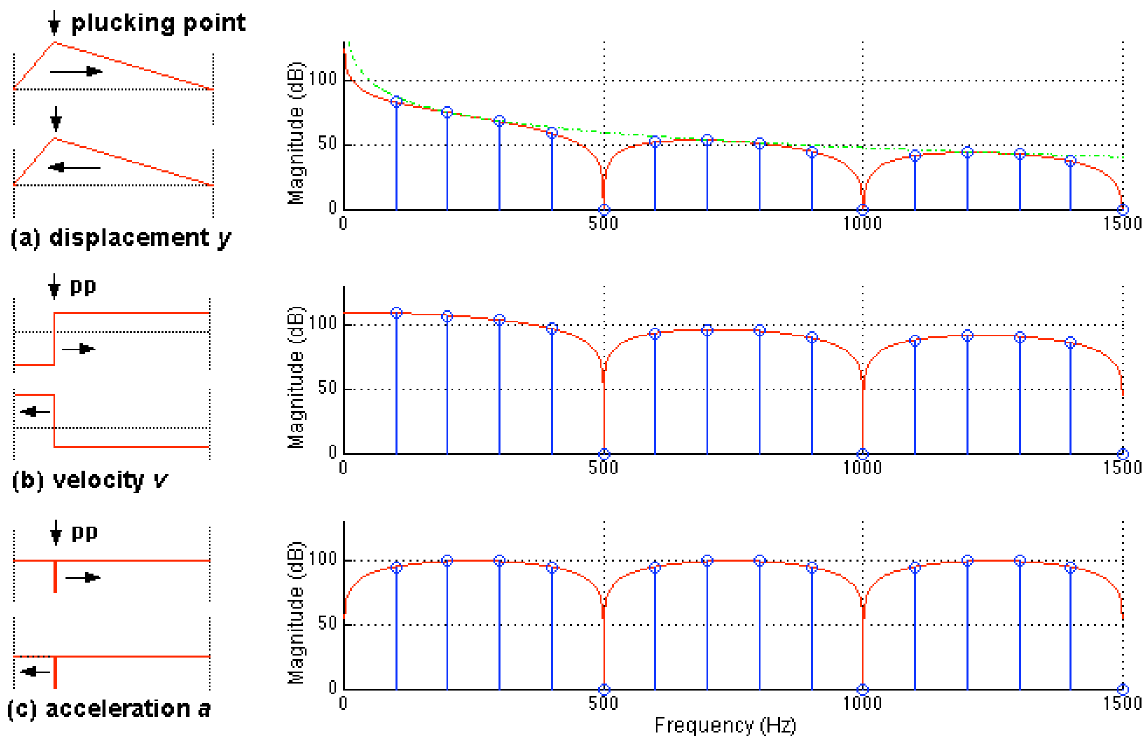


Fig. 4.6 Theoretical magnitude spectra for the displacement, velocity and acceleration variables. Fundamental frequency equals 100 Hz and relative plucking position is 1/5.

The equation equals 0 when $\frac{fR}{f_0}$ is an integer, that is when

$$f = \frac{nf_0}{R}.$$

The spectral envelopes for the different wave variables are displayed on Fig. 4.6. The factor $1/f^2$ in Eq. (4.23) is responsible for the -6 dB slope in the magnitude spectrum.

While there is no restriction on $1/R$ being an integer, the overall shape of the spectral envelope is dictated by $1/R$. For integer values of $1/R$, nulls occurs at harmonics which order is a multiple of $1/R$. For non-integer values of $1/R$, nulls in the spectral envelope occur at frequencies that are not necessarily related to the harmonic frequencies nf_0.

A typical magnitude spectrum is illustrated on Fig. 4.7 for a recorded guitar tone plucked 12 cm away from the bridge on a 58 cm open A-string (fundamental frequency = 110 Hz). The relative plucking position R is approximately 1/5 (12 cm / 58 cm = 1.483). If it were exactly 1/5 and if the string was ideal, all harmonics with indices that are multiples of 5 would be completely missing.

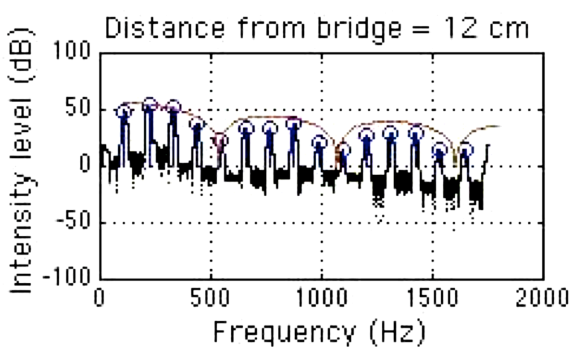


Fig. 4.7 Magnitude spectrum of a guitar tone and superimposed theoretical spectral envelope. A-string ($f_0 = 110$ Hz) is plucked at 12 cm from the bridge on a 58 cm string, resulting in a relative plucking position R close to 1/5.

4.4 Variation of brightness with plucking position

From the theoretical expression of the magnitude spectrum, spectrum-dependent perceptual measures can be derived, such as the spectral centroid which is correlated to *brightness* [88]. As guitarists intuitively associate increasing brightness with decreasing plucking distance from the bridge, we assume that it is possible to verify this correspondence by calculating

the spectral centroid of the power spectrum:

$$SC = \frac{\sum_{n=1}^{N} f_n C_v^2[n]}{\sum_{n=1}^{N} C_v^2[n]} \tag{4.24}$$

where $C_v[n]$ is the magnitude of the nth spectral component of the velocity wave (given by Eq. 4.21) and f_n its frequency .

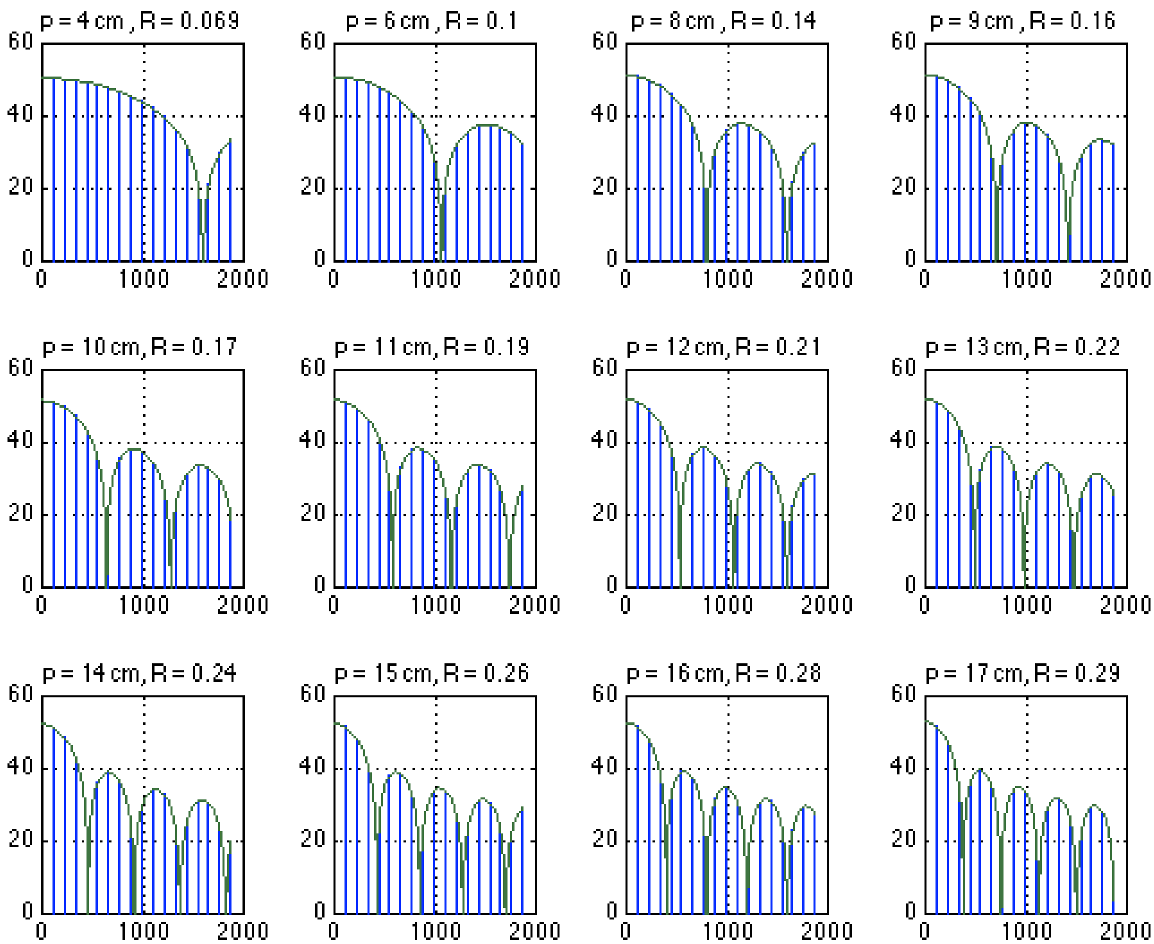


Fig. 4.8 Variation of the theoretical spectral envelope $C_v(f)$ (magnitude in dB vs frequency in Hz) with plucking position p ranging from 4 to 17 cm from the bridge.

Fig. 4.8 displays the plots of the theoretical spectra as for various plucking distances, calculated from the theoretical expression of the amplitude of the velocity modes. The velocity wave is considered here since pressure gradient microphones capture a wave analog

to velocity.

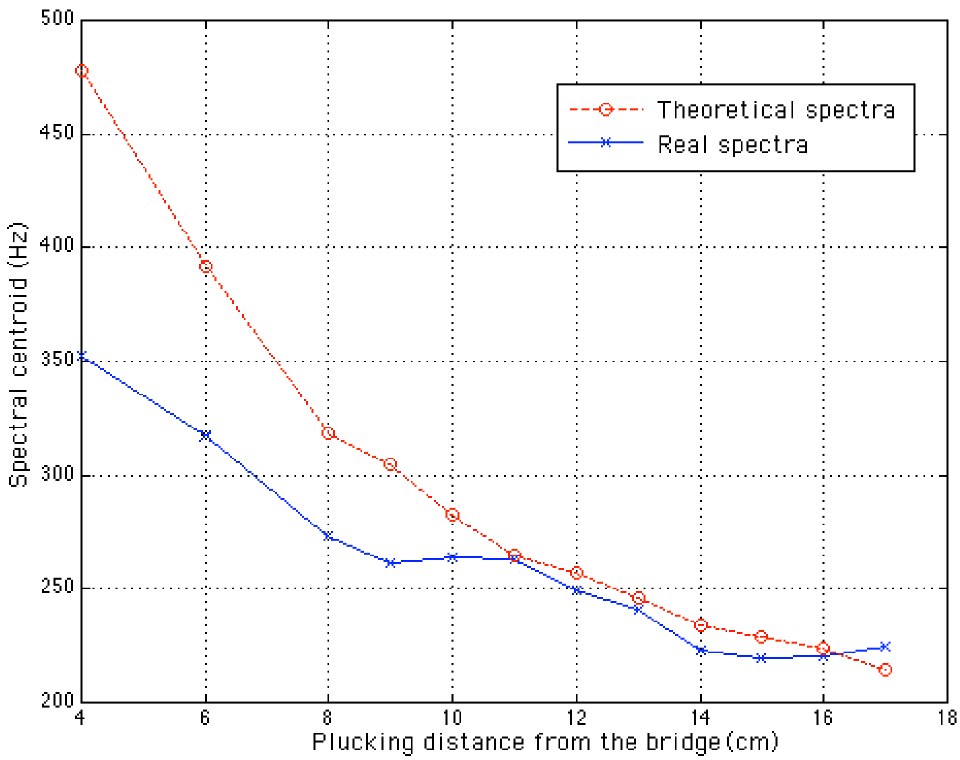


Fig. 4.9 Variation of the spectral centroid with plucking position p ranging from 4 to 17 cm from the bridge.

It is visually noticeable that the centre of gravity of the spectrum decreases as the plucking distance from the bridge increases. This trend is in fact confirmed by the plot displayed on Fig. 4.9, showing the spectral centroid of the theoretical spectra (shown on Fig. 4.8) as a function of plucking distance from the bridge. Also shown on Fig. 4.9 is the spectral centroid curve from the spectra of recorded guitar tones played with different plucking distances. The real data curve follows the same trend as the theoretical curve, although the spectral centroid is generally lower.

4.5 The real string

When compared to the ideal plucked string, a real guitar string reveals many differences. First of all, real strings have stiffness and damping, causing a lowpass filtering in the guitar tone spectrum. In addition, the lower guitar strings are inhomogeneous since they are

made from two different materials. Futhermore, the strings are mounted on the bridge of the guitar. Consequently, their vibration is influenced by the body modes of the top plate.

As illustrated on Fig. 4.10, the cutoff frequency in a guitar tone spectrum depends on two factors:

- the stiffness of the string,
- the width and the sharpness of the plectrum that excites the string.

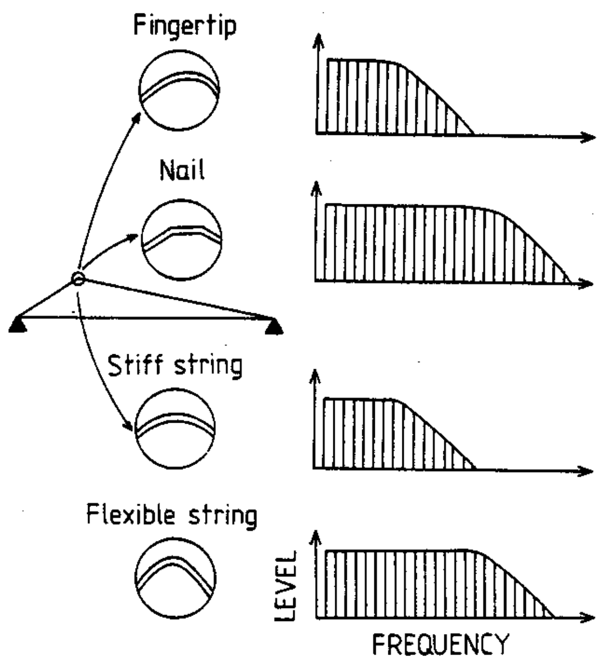


Fig. 4.10 Sketch of the influence of different ways of plucking and of stiffness of the strings on the resulting spectrum [10] (p. 16).

4.5.1 Partials are not completely absent in reality

Since real strings have stiffness and imperfections and since all plectra have finite width, it is far more accurate to state that the partials with nodes at the plucking position are strongly attenuated rather than completely absent.

4.5.2 Widening the excitation region

Widening the excitation region to an interval increases lowpass filtering of the excitation. For the sake of simplifying the model, the fact that the finger or plectrum exciting the string has a finite (non zero) touching width may be ignored. It is then assumed that the excitation acts at a single point.

4.5.3 Inharmonicity due to stiffness

Since the strings are under tension, they have some stiffness. To take into account the effect of the stiffness on the string motion, an extra term should be added to the wave equation Eq. (4.1) which becomes

$$\frac{\partial^2 y(x,t)}{\partial t^2} = \frac{T}{\mu}\frac{\partial^2 y(x,t)}{\partial x^2} - \frac{E\pi r^4}{\mu}\frac{\partial^4 y(x,t)}{\partial x^4} \tag{4.25}$$

where the string is assumed to be homogeneous (with constant mass per unit length μ), and where r is the radius of the cross sectional area, and E the Young's modulus of the string material [14].

The vibrational modes of a stiff string have frequencies which are not harmonically related and, in addition, the higher modes tend to be absent from the sound spectrum. In a simple model, one can imagine a stiff string exhibiting a rather smooth curve near the plucking position instead of a sharp angle, so that high frequency modes cannot be excited. This is illustrated on Fig. 4.10.

For the inharmonicity of the upper partial frequencies f_n, Morse [14] gives an approximate relation in the case of a stiff string as follows:

$$f_n = n f_1^0 \left(1 + \gamma + \gamma^2 + n^2 \frac{\pi^2\gamma^2}{8}\right), \tag{4.26}$$

where f_1^0 is the fundamental frequency of the same string without stiffness and where

$$\gamma = \frac{2r^2}{l}\sqrt{\frac{E\pi}{T}}$$

where r is the radius of the string, l is the length of the string, E is the Young modulus (proportionnal to stiffness), T is the tension.

The second and third terms in Eq. (4.26) show that the frequencies of the vibrational modes of a stiff string are higher than those of an ideal one given by Eq. (4.4).

The fourth term of Eq. (4.26), which contains the n^2 term, shows that this effect becomes more important as the frequency increases; the higher the frequency, the more it is shifted, and consequently the partials of a stiff string are not harmonically related [15]. Observing the formula, it can be concluded that the partials will be more harmonic if their order n is low (inharmonicity grows when going further away from the fundamental) and if the string is thin (small r), elastic (small E), long (great l) and tight (high tension T). Thicker strings are wound in order to reduce stiffness and consequently increase harmonicity.

4.5.4 String damping

On a vibrating guitar string, energy is lost through different mechanisms. As described by Fletcher [5], the main loss mechanisms are

- the internal damping of the string,
- the damping from the surrounding air,
- the transfer of energy to the guitar body through the moving ends of the string.

The internal damping is an inherent property of the material, independent of the string dimensions and tension. It is generally negligible for solid metal strings but may become the prime damping mechanism for gut or nylon strings, or for strings of nylon wound with metal.

The air damping is caused by the viscous flow of air around the moving string. It depends on the string radius and the frequency of oscillation in such a way that the high frequency modes of the string decay more quickly than the low frequency ones. Due to the air damping, the amplitude of vibration at a single frequency decays exponentially with time. In order to minimize the effects of air damping, a thick wire of dense material should be used.

The effect of energy transfer from the string to the guitar body depends on the properties of the string end supports. The frequencies of the vibrational modes of the string are lowered or raised, depending on the kind of support, and their decay time is affected.

The end support is characterized by its mechanical impedance, defined as the ratio between the applied force and the velocity of the support (F/v). Gough describes two

different types of end supports: the mass-like support and the spring-like support [4]. If the end support acts as if there were a mass connected to the string end, the motion of the support lags behind the driving force the string exerts on it. In this case, the node of string vibration is not created on the support itself, but on the string a short distance from the end; the wavelength of the string mode decreases and consequently the frequency increases. If the end support acts as if there were a spring connected to the string end, the node of vibration is created somewhere beyond the string support, inducing an increased wavelength and a decreased frequency. Generally, the mechanical impedance of the support changes with the frequency so that different frequencies are shifted by a different amount [15].

Chapter 5

The Plucking Effect as Comb Filtering

Contents

In this chapter, we explicitly represent the plucking effect. We derive a digital signal processing interpretation of the plucking string physical model which is a comb filter with delay $D = R/f_0$ (relative plucking position over fundamental frequency of the string). Then, the notion of "comb filter formant" is introduced. We also describe how to improve the control of waveguide-based synthesis of a plucked string which includes a comb filter to simulate the localized plucking excitation. We explain how the comb filter delay should be set for a realistic reproduction of the performance.

5.1 Digital signal processing interpretation of the plucked string physical model

In this section, we present a digital signal processing interpretation of the physical model of the plucked string derived in Chapter 4. The amplitude of the spectral components of the acceleration wave is given by

$$C_a[n] = 4K(R)f_0^2|\sin(n\pi R)| \tag{5.1}$$

In a simple digital physical model of a plucked-string instrument, the resonant modes translate into an all-pole structure (i.e. the harmonic structure of the signal), while the initial conditions (a triangular shape for the string and a zero-velocity at all points) result in a non-recursive FIR comb filter structure of the type

$$y[n] = x[n] - x[n-d] \tag{5.2}$$

where d is the delay expressed in number of samples. This comb filter constitutes the spectral envelope structure of the signal.

Eq. (5.2) is adequate for the digital interpretation of the acceleration variable along a plucked string since the acceleration impulse reflects negatively off the bridge, as illustrated on Fig. 5.1. Taking the z-transform of Eq. (5.2), we obtain

$$Y(z) = X(z) - X(z)z^{-d} = X(z)(1 - z^{-d})$$

from which we get the transfer function

$$H_d(z) = \frac{Y(z)}{X(z)} = 1 - z^{-d}.$$

Then we determine the magnitude response of that filter

$$\begin{aligned}
|H_d(e^{j\Omega})|^2 &= H_d(e^{j\Omega})H_d(e^{-j\Omega}) \\
&= \left(1 - e^{-j\Omega d}\right)\left(1 - e^{j\Omega d}\right) \\
&= 2(1 - \cos(\Omega d)) \\
&= 4\sin^2(\Omega d/2)
\end{aligned}$$

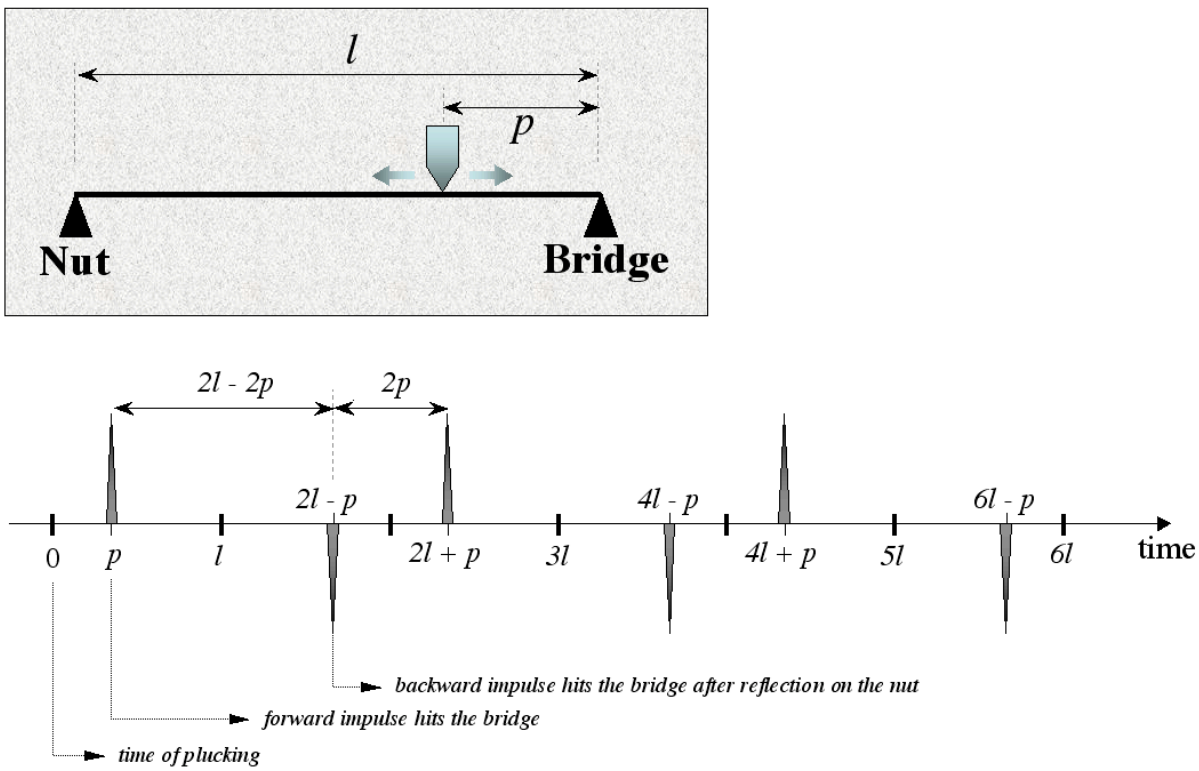


Fig. 5.1 Acceleration impulses received at the bridge after multiple reflections on the bridge and on the nut.

Hence, at a sampling rate f_s, the magnitude of the frequency response of this comb filter is given by

$$|H_d(e^{j\Omega})| = 2|\sin(\Omega d/2)| = 2|\sin(\pi\, d\, f/f_s)| \tag{5.3}$$

where the delay d can be a non-integer number of samples, corresponding to the time the wave needs to travel from the plucking point to the fixed end of the string (the bridge or the nut) and back ($2p$) as illustrated on Fig. 5.1. The magnitude response of the FIR comb filter for a 10 ms delay is shown on Fig. 5.2. As the fundamental period T_o corresponds to

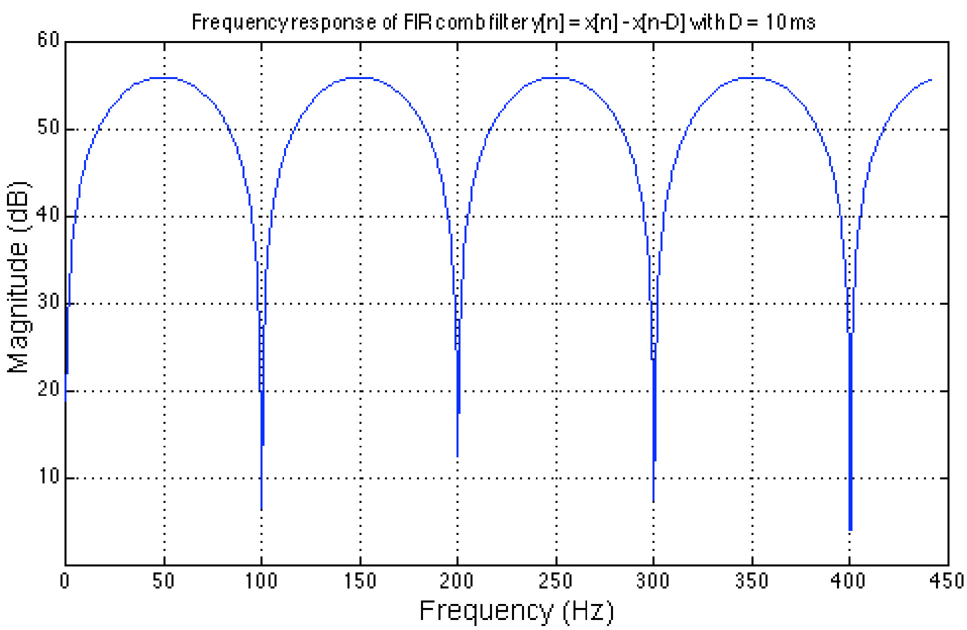


Fig. 5.2 Frequency response of FIR comb filter with a delay of 441 samples (10 ms).

the time the wave needs to travel along a distance that is two times the vibrating length of the string ($2l$), the relation between the comb filter delay D and the relative plucking position R is :

$$\frac{D}{T_0} = \frac{2p}{2l} = R \tag{5.4}$$

where $D = d/f_s$ is the delay expressed in seconds.

This relationship between the comb filter delay D and the relative plucking position R is the basis of the analogy between the physical model (Eq. 5.1) and its digital signal processing interpretation (Eq. 5.3). In fact, it is possible to verify that the arguments of

the sine functions in Eq. (5.3) and (5.1) are equivalent:

$$\pi\, d\, f/f_s = \pi D f = \pi R T_0 f = \pi R(f/f_0) = n\pi R \tag{5.5}$$

5.2 Comb filter formants

The notches in the magnitude spectrum of a FIR substractive comb filter occur at frequencies of components which, after being delayed, are still in phase with the original signal. In other words, the period of the component equals the delay or a submultiple of the delay. Hence, the notches occur at integer multiples of the inverse of the delay ($1/D$). The maxima, halfway between the notches, occur at odd integer multiples of half the inverse of the delay ($1/2D$).

The relationship between relative plucking position R and comb filter delay D is deducted from Eq. (5.4):

$$D = R\,T_0 = \frac{R}{f_0} \tag{5.6}$$

5.2.1 Comb filter formant central frequencies

Considering a string of length l plucked at a distance p from the bridge and resonating at fundamental frequency f_0, the frequency F_1 of the first local maximum in the comb-filter shaped magnitude spectrum equals the inverse of twice the delay D:

$$F_1 = \frac{1}{2D} = \frac{1}{2RT_0} = \frac{f_0}{2R} = \frac{l f_0}{2p} \tag{5.7}$$

The other local maxima ($F_2, F_3, ...$) in the magnitude spectrum are odd integer multiples of F_1. Since the comb filter peaks located at these frequencies F_n may act as formants, we will call them *comb filter formants*. Here we consider the literal definition of a formant: a frequency range in which amplitudes of spectral components are enhanced. In most cases, formant regions are due to resonances but in the present case, the local maxima do not correspond to resonances per se but rather to anti-notches.

Fig. 5.3 illustrates the case of a fundamental frequency f_0 equal to 100 Hz and a relative plucking position R equal to 1/5. The zeroes in the magnitude spectrum occur at integer multiples of $f_0/R = 500$ Hz and the local maxima occur at odd integer multiples of $f_0/2R =$

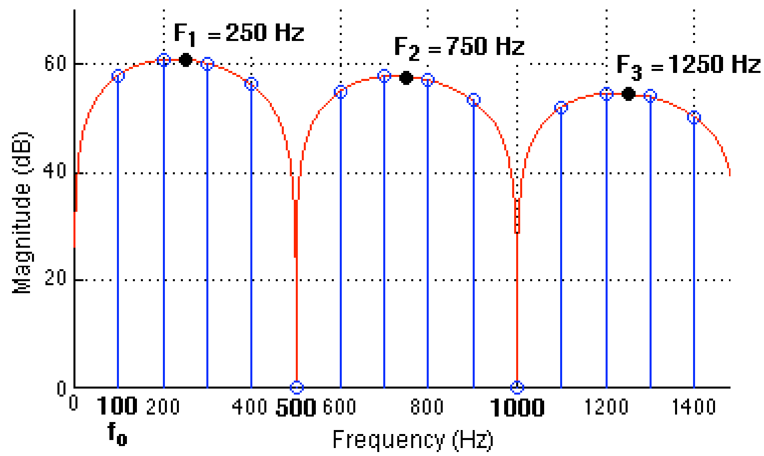


Fig. 5.3 Magnitude spectrum of the comb filter corresponding to a fundamental frequency of 100 Hz and relative plucking position of 1/5. Zeroes occur at integer multiples of 500 Hz and local maxima occur at odd integer multiples of 250 Hz.

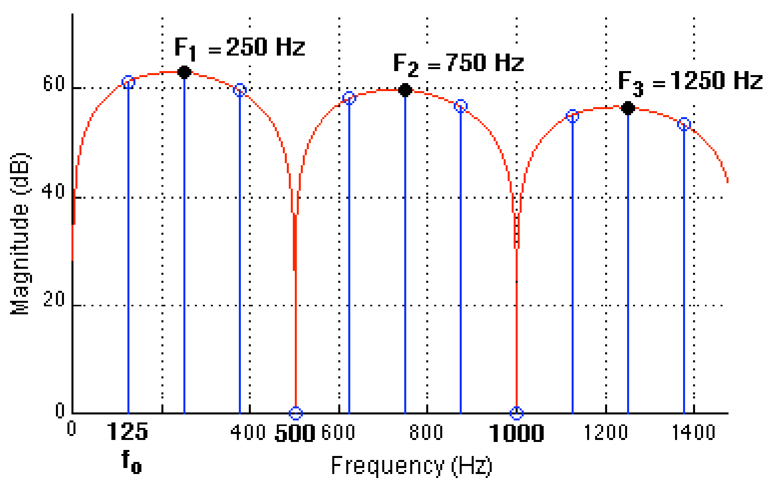


Fig. 5.4 Magnitude spectrum of the comb filter corresponding to a fundamental frequency of 125 Hz (a major third higher than the case illustrated on Fig. 5.3) and relative plucking position of 1/4. Zeroes occur at integer multiples of 500 Hz and local maxima occur at odd integer multiples of 250 Hz.

250 Hz. The frequencies $(F_1, F_2, F_3, ...)$ can be seen as the central frequencies of the comb filter formants.

It is interesting to note that the comb filter formant frequencies F_n are constant for a given absolute plucking position p on a given string, regardless of the note being played. For example, in order to play a note that is a major third higher than the note generated by an open string, the vibrating length of the string is shortened by a (5:4) factor by pressing the string with a finger against the corresponding fret. More generally, calling α the transposition ratio, the fundamental frequency f_0 is multiplied by the ratio α while the string length l is divided by the same ratio α (since f_0 is inversely proportional to the speed of sound on the string), hence

$$F_1 = \frac{(l/\alpha) \times (\alpha f_0)}{2p} = \frac{l f_0}{2p} = \frac{f_0}{2R}$$

By a simple inspection of Eq. (5.7) giving F_1 as a function of f_0 and l, one can see that the α's cancel each other. This is consistent with the fact that the product $l f_0$ is a constant for a given string and equals half the speed of sound c (as defined in Eq. (4.2)):

$$l f_0 = c/2 \tag{5.8}$$

It can be concluded that the comb filter formant frequencies on a given string occur at odd multiples of

$$F_1 = \frac{c/2}{2p} = \frac{c}{4p} \tag{5.9}$$

where p is the absolute plucking position and c is the speed of sound on the string.

Here is an example illustrating the fact that the comb filter formant frequencies are fixed for a given absolute plucking position p on a given string; if a 60 cm long open string tuned at 100 Hz is plucked at 12 cm from the bridge,

$$R = \frac{p}{l} = \frac{12}{60} = \frac{1}{5}$$

and

$$F_1 = \frac{f_0}{2R} = \frac{100}{2/5} = 250\,\text{Hz}$$

This case is illustrated on Fig. 5.3. Now, if the string is fingered to play a note a third

higher while the absolute plucking position is maintained, the different parameters become

$$l = \frac{60}{5/4} = 48\,\text{cm}, \quad p = 12\,\text{cm}, \quad R = \frac{p}{l} = \frac{12}{48} = \frac{1}{4}, \quad f_0 = 100 \times \frac{5}{4} = 125\,\text{Hz}$$

and

$$F_1 = \frac{f_0}{2R} = \frac{125}{2/4} = 250\,\text{Hz}$$

which is the same frequency as previously found (cf. Fig. 5.4).

It can also be shown that the cases R and $R' = 1 - R$ are equivalent since

$$|sin(n\pi(1-R))| = |sin(n\pi - n\pi R)| = |sin(n\pi R)|$$

for any integer n. For example, if the string is fingered in such a way that the vibrating length is 40 cm, plucking 10 cm from the bridge ($R = 10/40 = 1/4$) gives the same magnitude spectrum as plucking 30 cm from the bridge ($R = 30/40 = 3/4$).

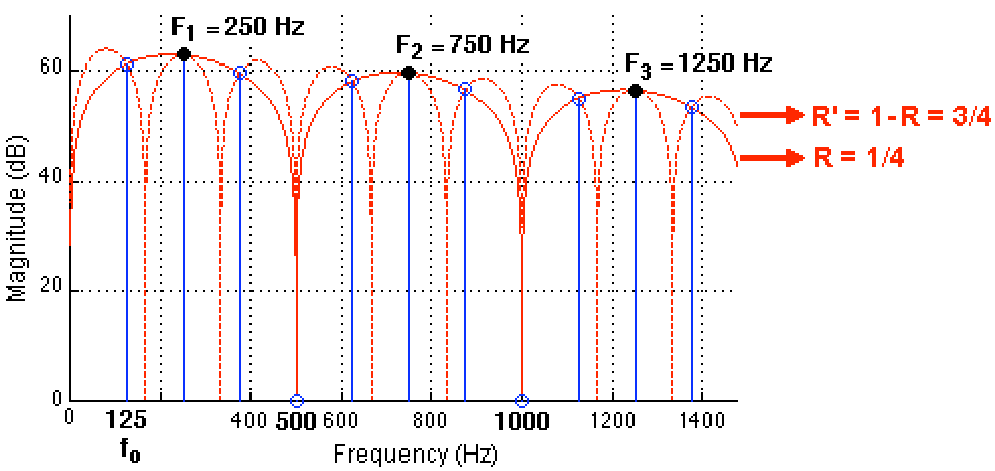


Fig. 5.5 Magnitude spectrum of comb filters with $f_0 = 125$ Hz and $R = 1/4$ or $3/4$.

With Eq. (5.7), one can determine the usual ranges of frequencies for the first formant frequencies F_1 of the different strings of a guitar. We consider a range of absolute plucking position going from 3 cm to 30 cm from the bridge on 60 cm strings tuned with the standard tuning. The smallest value for F_1 is obtained when plucking at the midpoint (30 cm) of

the open string:

$$F_1 = \frac{lf_0}{2p} = \frac{60 \times f_0}{2 \times 30} = f_0$$

The greatest value of F_1 is obtained when plucking very close to the bridge, say at 3 cm from it:

$$F_1 = \frac{lf_0}{2p} = \frac{60 \times f_0}{2 \times 3} = 10f_0$$

Therefore, the range for F_1 goes from f_0 to $10f_0$.

String number	Note name	Standard tuning frequency (f_0)	Range for the first formant frequency (F_1)
6	E (Mi_2)	83 Hz	[83 $\rightarrow$ 830] Hz
5	A (La_2)	110 Hz	[110 $\rightarrow$ 1100] Hz
4	D ($Ré_3$)	146 Hz	[146 $\rightarrow$ 1460] Hz
3	G (Sol_3)	202 Hz	[202 $\rightarrow$ 2020] Hz
2	B (Si_3)	248 Hz	[248 $\rightarrow$ 2480] Hz
1	E (Mi_4)	330 Hz	[330 $\rightarrow$ 3300] Hz

Table 5.1 Ranges for the first comb filter formant frequency for the six guitar strings (from 3 cm to 30 cm from the bridge).

Example: if the second string ($f_0 = 248$ Hz) is plucked at 15 cm from the bridge (assuming the string open and 60 cm long), the first formant frequency is

$$F_1 = \frac{lf_0}{2p} = \frac{60 \times 248}{2 \times 15} = 496 \approx 500 \text{ Hz}$$

The second resonance is centered approximately on $3 \times 500 = 1500$ Hz, the third on 2500 Hz, and so on.

5.2.2 Comb filter formant bandwidth

Having discussed the comb filter resonances as formant regions, we now specify the bandwidth of those formants. The magnitude spectrum being proportional to a sine function, we conclude that the 3 dB-bandwidth is given by

$$BW = \frac{F_1}{2} \tag{5.10}$$

which is the frequency range corresponding to a $[\pi/4, 3\pi/4]$ phase range and is the same for all comb filter formants.

Referring to the previously mentioned example illustrated on Fig. 5.3, the formant central frequencies are odd multiples of 500 Hz and the bandwidth of all formants is 500/2 = 250 Hz. Note that this is wider than the bandwidth of a usual vowel formant.

5.3 Digital modelling of plucked strings

Digital waveguide synthesis models are computational physical models which are made up of delay lines, digital filters, and often nonlinear elements. Waveguide-based digital models of plucked strings are described in [42], [38], [45], [39], [40] and [46].

In this section, we explain how the delay of the comb filter that simulates the localized plucking excitation ought to be set for a realistic reproduction of the performance.

5.3.1 Waveguide model of plucked strings

In a simple waveguide model (as described in [46]), the string is modelled with a dual delay-line as shown in Fig. 5.6. The total number of samples in the whole loop L corresponds

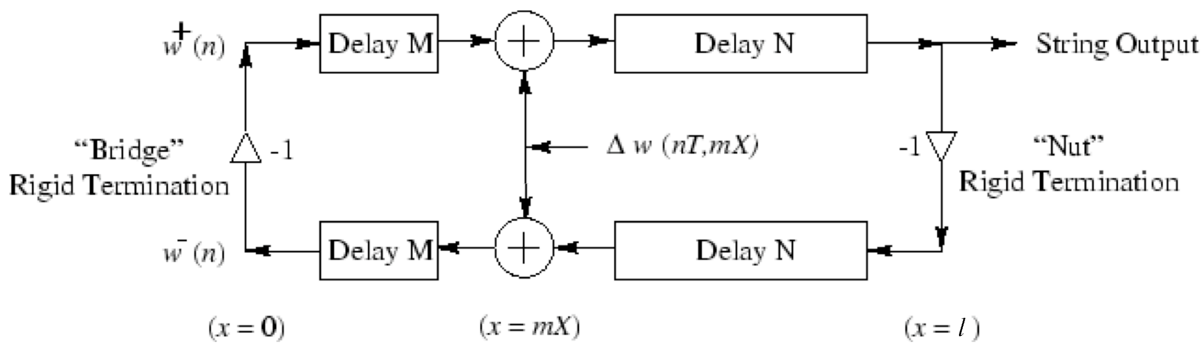


Fig. 5.6 Dual delay-line model for a guitar string [46].

to the fundamental period T_0 of the string. Hence, the delay L may be obtained from the ratio of the sampling frequency f_s over the fundamental frequency f_0:

$$L = \frac{T_0}{T_s} = \frac{f_s}{f_0} \tag{5.11}$$

The excitation $\Delta w(nT, mX)$ is introduced at M samples from the bridge termination. The delay L is then split in $L = 2M + 2N$ as illustrated in Fig. 5.6.

As the two effects of this dual delay-line model is an all-pole filter $H_L(z)$ that controls the modes of oscillation, and a comb filter $H_D(z)$ that controls the spectral envelope, the system can be split into a cascade of an all-zero filter and an all-pole filter, as in Fig. 5.7, where the comb filter delay $D = 2M$ samples, and the delay of the feedback loop $L = 2M + 2N$. Hence, the z-transform of the plucked string model is

$$H_S(z) = H_D(z)H_L(z) \tag{5.12}$$

The consolidation of the delays into a single delay-line string loop leads to a more computationally efficient synthesis model.

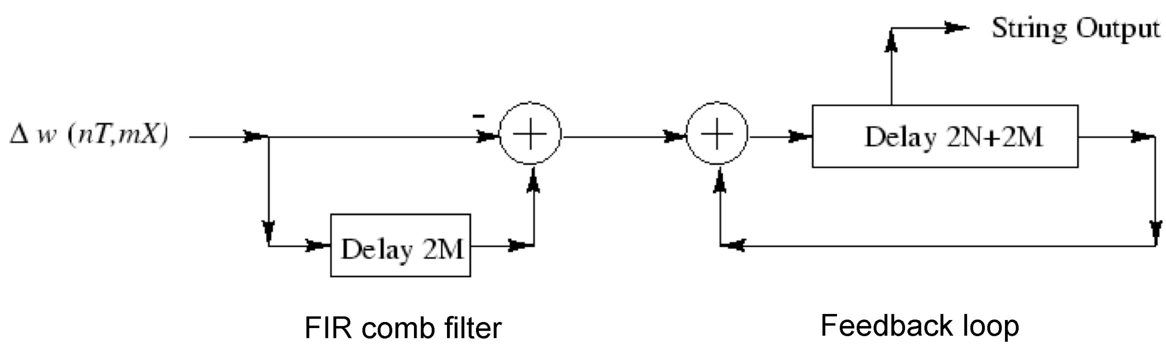


Fig. 5.7 Single delay-line modelling the string and factored out comb filter modelling the plucking effect [46].

The time-domain equation representing a single delay-line loop is

$$y[n] = x[n] + y[n - L] \tag{5.13}$$

This string model is a recursive comb filter which, by definition, has an infinite impulse response (IIR). Taking the z-transform, one obtains

$$Y(z) = X(z) + z^{-L}Y(z) \tag{5.14}$$

from which the z-transform of the transfer function is derived:

$$H_L(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - z^{-L}} \tag{5.15}$$

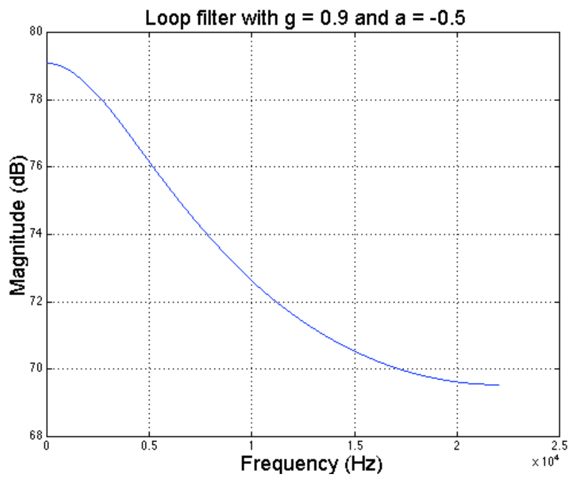


Fig. 5.8 Frequency response of a lowpass filter with gain $g = 0.9$ and $a = -0.5$.

To take into account the frequency-dependant damping, a low-pass filter is introduced in the feedback loop and the transfer function of the string model becomes

$$H_L(z) = \frac{1}{1 - H_{damp}(z)z^{-L}} \tag{5.16}$$

The lowpass filter $H_{damp}(z)$ is usually of the form

$$H_{damp}(z) = g\frac{1 + a}{1 + az^{-1}} \tag{5.17}$$

where g is a positive number slightly smaller than 1 and a is a small negative number between 0 and -1 [55].

In a single delay-line model, the plucking point equalizer consists in a comb filter with z-transform

$$H_D(z) = 1 - z^{-D} \tag{5.18}$$

On Fig. 5.9 are plotted the frequency responses of the string model and of the plucking

point equalizer separately (top) and combined (bottom).

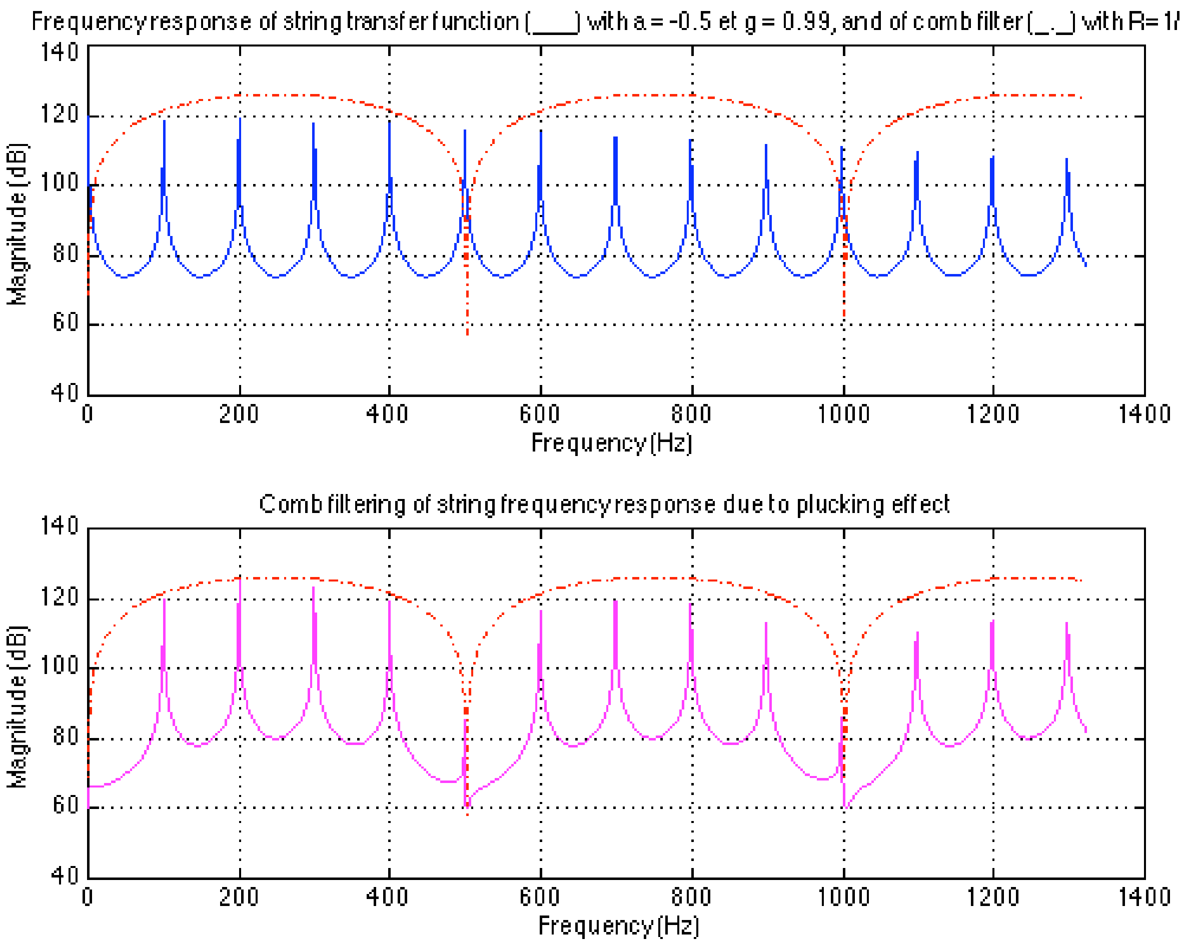


Fig. 5.9 In top figure, the magnitude response of the feedback loop $|H_L(e^{j\Omega})|$ is displayed. In bottom figure, the global magnitude response $|H_S(e^{j\Omega})|$ is displayed, including the effect of the comb filter. The magnitude response of the comb filter $|H_D(e^{j\Omega})|$ is superimposed on both figures, traced with a dotted line.

5.3.2 Controlling the comb filter in an realistic manner

When digital waveguide models are used to synthesize a guitar piece, the value of the comb filter is often set to a constant value. As seen in the previous section, the comb filter delay

depends of the relative plucking position R:

$$D = \frac{R}{f_0} = \frac{p}{l f_0} \tag{5.19}$$

We also saw that the comb formant frequency F_1 is constant on a given string and for a given plucking position p. From Eq. 5.9, we obtain

$$D = \frac{1}{2F_1} = \frac{2p}{c} \tag{5.20}$$

where p is the absolute plucking position from the bridge and c is the speed of sound on the string. Using Eq. 5.20, the comb filter delay is calculated from the fingering information provided by an experienced guitarist or calculated by an automatic fingering generator (e.g. prototype described in [44]). The choice of string determines c and the plucking position determines p. As illustrated in Table 5.2, the delay varies greatly for all 6 strings plucked at a given absolute plucking position from the bridge. Knowing the angle of the hand and forearm with respect to the strings axis, the difference in distance from the bridge for the different fingers of the hand can be modelled. The pluck by the index finger is generally slightly further away from the bridge than the pluck by the ring finger.

String number	Note name	Standard tuning frequency (f_0)	Delay (in ms) for $p = 12$ cm	Delay (in samples) for $p = 12$ cm
6	E (Mi_2)	83 Hz	2.4	106
5	A (La_2)	110 Hz	1.8	80
4	D ($\mathrm{Ré}_3$)	146 Hz	1.4	60
3	G (Sol_3)	202 Hz	1.0	44
2	B (Si_3)	248 Hz	0.8	36
1	E (Mi_4)	330 Hz	0.6	27

Table 5.2 Value of the comb filter delay in ms and in samples (at $f_s = 44100$ Hz) for each of the 6 strings of a guitar plucked 12 cm from the bridge (strings are 60 cm long).

Part II

Guitar Timbre Perception

Chapter 6

Timbre: a Multidimensional Sensation

It is the immense difference between the physical acoustic signal on the one hand and the perceptual-cognitive world on the other hand that has frustrated theorists and researchers.
Stephen Handel [89] (p. 265).

Contents

A very general definition of timbre is given by the American Standard Association: "timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" [85]. Actually, timbre can be studied at different levels. From a *macroscopic* point of view, one may examine the differences between the timbre of a violin and the timbre of a guitar. From a *microscopic* point of view, one may examine the differences within these instrumental categories, such as subtleties between a Stradivarius and a Guarnerius violin, or a Ramirez and a Rubio guitar. Further and even more important from the performer's point of view, one can further examine the timbral difference between a note played ponticello (close to the bridge) and tasto (close to the nut) on the same instrument (see Chapter 3).

Before exploring the vocabulary used by guitarists to describe guitar tones, the main theories of timbre perception and methods used to study the perception and the description of timbre are presented.

6.1 The parameters of timbre

J. F. Schouten [96] defines five parameters of timbre:

- the temporal envelope in terms of rise time, duration and decay;
- the prefix, which is the onset of a sound, quite dissimilar to the ensuing lasting vibration;

- the spectral envelope (amplitude profile of the partials of a sound);
- the change of spectral envelope (formant glide) and fundamental frequency (micro-intonation);
- the ratio between harmonic and noise-like character (the scale ranging from perfect harmonicity through pseudo-harmonicity to random noise).

According to G. von Bismarck [115], steady speech sounds and musical sounds differ mainly in the following physical parameters:

- the frequency location of the whole spectrum;
- the slope of the spectral envelope;
- the frequency location of energy concentrations (e.g. formants) within the spectrum.

In the case of plucked string instruments, many of the timbral parameters are interdependent since the manipulation of the string produces almost all of the aforementioned changes in timbre.

6.2 The description of timbre

The tone-qualities of instruments may be described and compared in a number of ways.

6.2.1 The source-mode of timbre perception

Timbre can be described in terms of the structural characteristics of the instruments producing the tones. For example, string and wind instruments may conveniently be separated by virtue of the fact that one group uses a vibrating string, whereas the other uses a vibrating wind column. The former further subdivides into instruments whose string is set into vibration by a bowing motion, a pluck, or a striking excitation. Handel proposes an explanation for timbre perception, stating that the subjective identification of timbre could involve the observer's perception of the physical mechanisms and actions in the sound production [89]. This is the source-mode of timbre perception, as opposed to the interpretative mode of timbre perception [88]. It is also interesting to realize that the mechanics and the materials of vibrating systems are the basis for traditional Western musical instruments, as

well as World instrument classification systems (e.g. von Hornbostel & Sachs classification in aerophones, chordophones, membranophones and idiophones [160]).

6.2.2 Harmonic theory of tone-quality

The harmonic theory of tone-quality states that "[...] all varieties of tone quality are due to particular combinations of a larger or smaller number of simple tones" [13]. Helmholtz completes this statement with: "quality of a musical tone depends solely on the number and relative strength of its partial simple tones [...]" [90].

Helmholtz was the first to attempt to find acoustic correlatives for descriptive qualitative terms such as pleasant, harmonious, rich, poor, hollow, etc. With resonators of different sizes, Helmholtz analyzed the tones of some instruments and attributed tone-quality to the relative strengths of the overtones, independent of the fundamental [90]. He distinguishes different types of tone-colour, based on the presence or absence of higher partials:

- simple tones, such as tuning forks and wide stopped organ-pipes, have a very soft, pleasant sound, free from all roughness; when low in pitch, there are dull;
- tones in which the first 6 partials are moderately loud, such as those of the piano, the open organ-pipe, and the French horn, are more harmonious and musical than simple tones; they are rich and splendid compared to simple tones and are sweet and soft when the partials higher than the 6th are absent;
- when only odd-numbered partials are present (as is, to a large extent, the case with the clarinet), the tone is hollow;
- if many predominating high partials (above the 6th partial) are present, the tone is nasal;
- if partials higher than the 6th or the 7th are distinct, the tone is cutting and rough, as is the case with the bassoon, oboe and brass instruments;
- if the fundamental predominates, the tone is rich;
- if the fundamental does not predominate, the tone is poor.

6.2.3 Formant theory of tone-quality

Thirty years after the publication of *The Sensations of Tone* by Helmholtz, Erich Hermann-Goldap [91] published a paper in which he contradicts Helmholtz with his *formant theory of tone-colour*, explaining that each tone-colour has a characteristic range of overtone strength which does not vary with the fundamental pitch. As does Helmholtz, Hermann-Goldap subjects the graphic representations of the vibrations of the instruments to Fourier analysis and makes various observations:

- the horn has a second formant, which appears when it is played loudly;
- the formants of the oboe, flute, clarinet, and trumpet lie in the same register as those of the trombone and horn;
- the instruments cannot be distinguished solely on the basis of the formants' position; one must also consider the amplitude of the fundamental:
  - if the amplitude of the fundamental is small when compared to that of the formant, the tone is <u>sharp</u> (cf. the oboe and the trumpet);
  - as the amplitude of the fundamental approaches that of the formant, the tone becomes more <u>full</u> and more <u>pleasant</u> (cf. the horn and the softly-played trombone);
  - when the amplitude of the fundamental surpasses that of the formant, the tone first become <u>soft</u> and then becomes <u>nasal</u>.

Youngblood [158] regards the human voice as the most remarkable of all musical instruments, every vowel being a different tone-colour. He ponders that "it is therefore difficult to understand why one would describe the timbres of man-made instruments in terms of a harmonic theory and those of the natural instrument in terms of a formant theory. If tone-quality be the key issue, then it seems that the same theory should apply to both." [158] (p. 57). In other words, Youngblood suggests that the formant theory is more appropriate since there is no reason to consider instrumental timbre and vocal timbre differently.

Indeed, in many cases, a fixed formant structure gives a timbre that varies less with frequency than a fixed spectrum [95] (p. 115).

6.2.4 Verbal description of timbre by sound engineers

Sound engineers enhance the sound of a recording by boosting some frequency regions, according to the desired effect.

Boosting the signal in the *low bass* range (1st and 2d octaves, from 20 to 80 Hz) gives sound fullness and power. The *midrange* covers the 5th, 6th and 7th octaves (from 320 to 2560 Hz). For many sounds, the fundamental falls in the 5th octave. Boosting the 6th octave gives the sound a horn quality. Boosting the 7th octave gives the sound a tinny quality. The *upper midrange* is the 8th octave (from 2560 to 5120 Hz). Boosting this range improves intelligibility and adds presence to speech. Finally, boosting the *treble* range, covering the 9th and 10th octaves (from 5120 to 20000 Hz) adds sharpness and crispness to the sound.

6.3 Methods for studying timbre perception

6.3.1 Multidimensional scaling analysis

The Multidimensional Scaling (MDS) measurement method has been employed in an attempt to find the dimensions of timbre perception. This method is based on similarity judgements of sounds. The number of the resulting dimensions carrying nearly all the variance is generally much smaller than the number of signal variables.

Pols used MDS for a restricted set of vowel-like sounds and found that three orthogonal dimensions are nearly sufficient to describe the timbres [137]. Early studies on instrumental timbre were performed by David Wessel and John Grey in the late 1970's on a data set of 16 instrumental timbres [87]. As shown on Fig. 6.1, timbral features such as brightness (associated with the spectral centre of gravity), spectral flux and transients density were identified. It is important to note that these axes were used to differentiate between different orchestral instruments – a macroscopic view of timbre – as opposed to differentiating between the possible palette of timbres in a single instrument – a microscopic view of timbre.

6.3.2 Semantic differential method

With the Semantic Differential (SD) method, sounds are rated on many category scales, the endpoints of which are characterized by opposite verbal attributes such as sharp-dull,

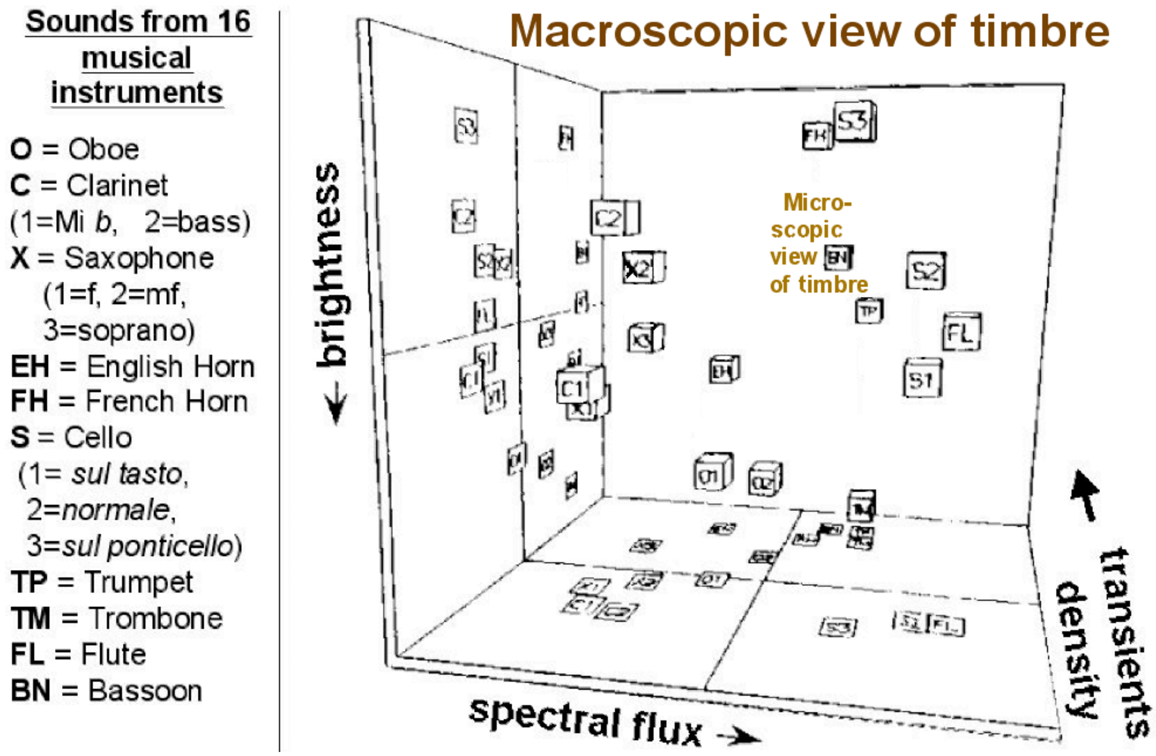


Fig. 6.1 Macroscopic and microscopic views of timbre (after [87]).

rough-smooth or concentrated-diffuse. This method, with subsequent factor analysis, was applied by G. von Bismarck to the perception of the timbre of complex steady sounds (equalized in loudness, pitch and duration) [115]. The experiment was designed to test the following hypothesis: timbres of sounds can be uniquely described if the sounds are rated on a few scales which are characterized by verbal attributes.

6.3.3 Free verbalization method

Critics of verbal scaling methods justifiably noted that the pre-selection of verbal attributes which characterize SD-scales may strongly affect the results, since these scales do not necessarily conform with those a subject would use spontaneously [92]. The pre-selected scales may omit important aspects of timbre and may contain irrelevant scales. To address this problem, instead of asking to rate a sound quality according to predefined scales, the experiment consists of collecting spontaneous comments on sounds. This method based on free verbalization were used namely by Faure [101] and Samoylenko et al. [114].

6.3.4 Questionnaire-based method

In a study on the verbal description of timbre in Czech language, Moravec & al. [110] submitted questionnaires to musicians. The participants were asked to write down in free order the words and expressions which they use for the description of timbre, as well as groups of synonyms and antonyms.

6.4 Some results from studies on timbre perception

6.4.1 Bismark scales

Among the early studies on the verbal descriptors of timbre, the study by G. von Bismark [115] is probably the most precise. In this study, pairs of opposite attributes, such as dark-bright or smooth/rough, characterized the endpoints of scales, on which 35 sounds were rated by two groups of subjects possessing either intensive or no musical training. The sounds differed systematically in the parameters of the spectral envelope.

Factor analysis of the scale correlations provided four orthogonal factors which extracted 90 % of the variance. The factor carrying most of the variance (44 %) was represented by the scale dull-sharp. The scales representing the other factors appeared to be less suitable for the description of timbre in general than the scale dull-sharp. von Bismarck first surveyed studies in which scales were used (in particular studies by Solomon [112] and Kerrick et al. [104]) and found a total of 69 scales. Each of these was rated in turn for its suitability in describing timbre. From the 35 scales with the highest mean ratings, seven were eliminated because they were synonymous with other scales or had been proven to be unsuitable for the SD-analysis of timbre. Finally, von Bismarck selects 28 scales which were considered a representative sample:

weak-strong, gentle-violent, fine-coarse, reserved-obtrusive, dark-bright, dull-sharp, soft-hard, dim-brilliant, relaxed-tense, calm-restless, rounded-angular, dampened-ringing, smooth-rough, heavy-light, broad-narrow, wide-tight, thick-thin, clean-dirty, full-empty, solid-hollow, colourful-colourless, pure-mixed, simple-complex, compact-scattered, interesting-boring, lively-dead, pleasant-unpleasant, open-closed.

The degree of inter-individual scatter between ratings obtained with a particular scale was considered as a criterion for its psychophysical usability. The standard deviation of the

ratings was chosen as a measure of the scatter. Averaging the variances over all subjects yielded the following four scales with the smallest scatter in ascending order: round-angular, gentle-violent, reserved-obtrusive, soft-hard. The scales with the largest scatter were: full-empty, solid-hollow, colourful-colourless, open-closed, dead-alive, interesting -boring. von Bismarck concludes that these scales are not very useful for the measurement of timbre.

For musician and non-musician subjects, the total variance was almost completely extracted by three factors. A considerable portion of the variance was extracted by the first factor characterized by the attributes hard, angular, obtrusive, violent, sharp, rough, tense and unpleasant. The second factor was represented by the attributes ringing for both groups of subjects and by narrow for the musicians.

Varimax rotation of the factor axes led to the following interpretation: it appears that the timbre of the 35 sounds can be almost completely described if the sounds are rated on the largely independent scales : dull-sharp, compact-scattered, full-empty, colourful-colourless.

The ratings of both groups of subjects show that the sharpness of the sound increased when either the upper limiting frequency or the slope of the spectral envelope was raised. von Bismarck concluded that the attribute sharpness is primarily determined by the frequency position of the overall energy concentration of the spectrum rather than the shape of the spectral envelope. The attribute compactness was clearly used by both groups of subjects to differentiate between noise and tone stimuli.

Only the scales representing the first factor showed a small scatter of individual ratings. The author concludes that the only scales that are applicable scales for the measurement of timbre are dull-sharp, soft-hard and round-angular.

The general conclusion of von Bismark's study is that verbal attributes may be used in a consistent manner by subjects to describe different aspects of timbre (parameters of sound other than loudness and pitch). The majority of these attributes can be represented by the attribute sharpness, which is determined by the frequency location of the overall energy concentration of the spectrum. von Bismark further concludes that it does not seem possible to verbally describe in a psychophysically applicable manner other aspects of timbre not accounted for by sharpness [115].

6.4.2 Kendall & Carterette experiments: from bipolar to unipolar scales

The study by Kendall & Carterette [103] is similar to Bismark's study, although only 10 semantic scales were used to rate 10 sounds. The sounds were mixtures of two wind instruments (among the flute, clarinet, saxophone, trumpet and oboe). With bipolar scales (i.e. dull-sharp), the results were not conclusive. In a second experiment, applying a technique called *Verbal Attribute Magnitude Estimation* (VAME), unipolar scales were used (i.e. sharp-not sharp). Saxophone sounds were differentiated from the other sounds along the scales loud, heavy and hard. In a third experiment, the scales were modified in order to include more musical terms found in an orchestration treatise by Piston [111]. Principal component analysis of the results obtained for 21 semantic unipolar scales extracted 4 principal dimensions: power, strident, plangent and reed [88].

6.4.3 Verbal correlates of perceptual dimensions

A study conducted by Faure [101] sought to define the verbal correlates of perceptual dimensions observed in multidimensional studies on timbre. Musicians and non-musicians were asked to rate the dissemblance between pairs of timbres and then to freely describe all the similarities and dissimilarities between these timbres. Only some descriptors were correlated to one perceptual dimension at a time: the adjective dry was correlated to the *rise time of the temporal envelope*, the adjective round was correlated to the *spectrum central centroid* and the adjective bright was correlated to the *spectral flux*.

6.4.4 From a macroscopic to a microscopic view of timbre

Among studies aiming to identify the perceptual dimensions of timbre, a very small number investigate the timbre nuances of one particular instrument (i.e. the violin in [113] and the oboe in [102]). Similarly, very few studies explore the vocabulary used to describe the dimensions of one particular instrument's timbre space (i.e. the pipe organ in [99], the violin in [109], and the electric guitar in [106, 107]).

The next chapter reports the results of our study on the verbal descriptors for the timbre of the classical guitar.

Chapter 7

Verbal Descriptors for the Timbre of the Classical Guitar

It is interesting to note that, in two musical universes apparently distinct – oral traditions and contemporary creation –, the difficulty of verbalizing the mechanisms involved in perceiving, evaluating and producing music, contributes to reinforcing a widespread opinion according to which musical activity would not be as systematic as musicians maintain, in other words, that all attempts to model this activity would fail. Such a conception reveals on one hand a misunderstanding of the experimental scientific thinking process, where "failure" is a source of learning; on the other hand, it encourages, in traditional societies, the denial of highly sophisticated pedagogical methods to the profit of what can be called learning by imitation.

Bernard Bel [162] (p. 25)

Contents

The guitar is an instrument that gives the player great control over the timbre. Various plucking techniques involve varying the finger position along the string, the inclination between the finger and the string, the inclination between the hand and the string and the degree of relaxation of the plucking finger. Guitarists perceive subtle variations of these parameters and they have developed a very rich vocabulary to describe the brightness, the colour, the shape and the texture of the sounds they produce on their instrument. Dark, bright, chocolatey, transparent, muddy, wooly, glassy, buttery, and metallic are just a few of those adjectives. This chapter reports experiments and resulting data from a study based on the concepts and methodologies presented in Chapter 6, whose aims are to establish an inventory of adjectives used by guitarists to describe timbre and to investigate the correlations between plucking techniques and verbal descriptors.

7.1 An inventory of timbre descriptors for the classical guitar

7.1.1 Methodology

As a starting point for the exploration of the timbre space, we inquired about the timbre descriptors commonly used by professional musicians.

22 guitarists were asked to select 10 adjectives that best describe timbre nuances produced on their instrument (see Appendix C for questionnaire). A list of 50 adjectives was provided, but the participants were encouraged to use any term they deemed appropriate; this assured that the participants were only to define adjectives that were meaningful to them. They were to provide synonyms, antonyms and an English or French translation accordingly. After intuitively describing each timbre ("How does it sound?"), the participants

were asked to explain its corresponding gesture (“How is it executed ?”). The participants were classical guitar performance majors at the Université de Montréal; they had studied with different teachers upon entry into the programme. Most were francophones from Québec.

7.1.2 Classification of collected data

In total, the 22 guitarists defined about 80 different adjectives (see Appendix D). Some adjectives were chosen more often than others, as shown in Table 7.1. The adjectives metallic, round and bright were chosen by more than half the participants. The adjectives thin, warm, velvety, nasal and dry were chosen by about a third of the participants.

Number of definitions	Adjective in French	English translation
14 x	métallique	metallic
13 x	rond	round
12 x	brillant	bright
8 x	mince, chaleureux	thin, warm
7 x	velouté, nasillard, sec	velvety, nasal, dry
5 x	rugueux, sombre, sourd	rough, dark, muted
4 x	doux, épais, incisif, pulpeux, résonant	sweet, thick, sharp, pulpous, resonating
3 x	clair, creux, cuivré, lumineux, naturel, ouvert, plein, spongieux, transparent, voilé	clear, hollow, brassy, luminous, natural, open, full, spungy, transparent, veiled
2 x	étouffé, ovale, percussif	damped, oval, percussive

Table 7.1 Histogram for the adjectives which were defined by at least two participants. In the left column are given the numbers of participants (out of 22) who defined each adjective.

Additional adjectives were provided within definitions of the initially selected terms. Table 7.2 presents, in alphabetical order, the 108 adjectives compiled; this includes the ±80 adjectives that were initially selected (direct citations), as well as the ±30 that appeared within definitions (indirect citations). English translations are provided. Some translations were given by the participants in this questionnaire-based study. They were all checked and confirmed by our collaborator guitarist Peter McCutcheon who is perfectly bilingual.

French	English	French	English
001. Aiguisé	Sharp	055. Lyrique	Lyrical
002. Apaisant	Appeasing	056. Maigre	Skinny, meagre
003. Artificiel	Articificial	057. Martelé	Martelé
004. Basson	Bassoon	058. Mat	Dull / Mat / Matted
005. Brillant	Bright/Brilliant	059. Métallique	Metallic
006. Bruit blanc	White noise	060. Mielleux	Honeyed
007. Bulbeux	Bulbous	061. Mince	Thin
008. (Avec) caractère	With character	062. Mordant	Biting
009. Cassant	Brittle	063. Morne	Gloomy
010. Caverneux	Cavernous	064. Mou	Soft, limp
011. Chaleureux	Warm	065. Mouillé	Wet
012. Chaud	Warm	066. Mystique	Mystical
013. Chocolaté	Chocolatey	067. Nasillard	Nasal
014. Clair	Clear	068. Naturel	Natural
015. Clarinette	Clarinet	069. Nerveux	Nervous
016. Boîte à musique	Music box	070. Ouateux	Cottony
017. Collant	Sticky	071. Ouvert	Open
018. Confus	Muddled	072. Opaque	Opaque
019. Coulant	Flowing	073. Ovale	Oval
020. Coupant	Slicing	074. Perçant	Piercing
021. Crémeux	Creamy	075. Percussive	Percussive
022. Creux	Hollow	076. Pétillant	Sparkling
023. Criard	Shrill	077. Piquant	Pointed
024. Cristallin	Crystallin	078. Plat	Flat
025. Cuivré	Brassy	079. Plein	Full
026. Dense	Dense	080. Pleurnicheur	Whining
027. Doux	Sweet	081. Présent	Present
028. Dur	Harsh	082. Profond	Deep
029. Duveteux	Feathery, downy	083. Pulpeux	Pulpy
030. Éclatant	Bright, shining	084. Raboteux	Scraping
031. Émoussé	Blunt, dull	085. Rêche	Harsh
032. Entier	Whole	086. Résonant	Resonating
033. Enveloppant	Enveloping	087. Rêveur	Dreamy
034. Épais	Thick	088. Riant	Laughing
035. Estompé	Softened	089. Riche	Rich
036. Étouffé	Damped	090. Robuste	Robust
037. Explosif	Explosive	091. Rond	Round
038. Faible	Weak	092. Rugueux	Rough
039. Fermé	Closed	093. Sec	Dry

French	English	French	English
040. Feutré	Felty/Velvety	094. Sombre	Dark
041. Fibreux	Fibrous	095. Sourd	Matted/Surd
042. Flat	Flat	096. Spongieux	Spongy
043. Florissant	Blooming/Blossoming	097. Tight	Tight
044. Foncé	Dark	098. Terne	Colourless, drab
045. Fougueux	Wild	099. Tranchant	Slicing
046. Fracassant	Shattering	100. Transparent	Transparent
047. Gras	Fat	101. Transperçant	Piercing
048. Guimauve	Marshmellow	102. Vaporeux	Vaporous
049. Incisif	Incisive, sharp	103. Velouté	Velvety
050. Laiteux	Milky	104. Vif	Quick
051. Large	Large	105. (Avec) vitalité	Vivacious
052. Lisse	Smooth	106. Vitreux, vitré	Glassy
053. Lourd	Heavy	107. Voilé	Veiled
054. Lumineux	Luminous	108. Woody	Woody

Table 7.2 Adjectives qualifying timbre in French and English.

One participant spontaneously provided an annotated figure indicating on a guitar the different locations corresponding to different timbre descriptors (Fig. 7.1), thereby reasserting the important role of plucking position in timbral variations.

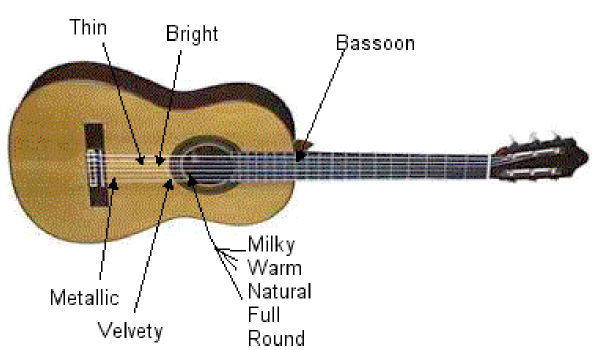


Fig. 7.1 Timbre descriptors and corresponding plucking locations along the string according to guitarist Zane Remenda (with permission).

Most guitarists have little knowledge about the acoustical and perceptual nature of sound and are not accustomed to describing their sounds in an objective and quantitative way. Therefore, most of the collected definitions contained analogies with other sounding objects or borrowed vocabulary from other sensory modalities. In the table displayed in Fig. 7.2, the adjectives are classified into different sensory categories.

- Category 1 - **Corps sonore**: refers to the sound of another sounding object (ex: bassoon)
- Category 2 – **Luminosité**: refers to the brightness/darkness associated with the sound (ex: shining)
- Category 3 – **Forme**: refers to the shape of the sound mental representation (ex: round)
- Category 4 – **Matière**: refers to a surface texture or a material property (ex: glassy)
- Category 5 – **Saveur**: refers to a food texture or flavour (ex: creamy)
- Category 6 – **Caractère**: emotion or character transmitted by the timbre (ex: vivacious)

Note that when two adjectives are listed in the table on the same line and separated by a slash bar, they are opposites of one another (example: Thin / Thick).

7.1.3 Organization of adjectives in clusters

In order to better organize the adjectives and specify their meaning, we have used the lists of synonyms provided by the participants in a direct or in an indirect way. In fact, in order to define an adjective, participants often referred to other timbre descriptors. Consequently, many synonyms were given in the definitions themselves. For example, participant #1 defined a <u>bright</u> sound as a "<u>clear</u> and <u>piercing</u> sound, sometimes <u>metallic</u> to a certain extent"; participant #8 wrote that a <u>bright</u> sound is "at mid-way between a <u>round</u> sound and a <u>metallic</u> sound. It is <u>clear</u>, <u>pure</u>, <u>franck</u> and <u>shining</u>".

In an attempt to establish a map for the mental representation of timbre according to this group of guitarists, we organized the adjectives into clusters, where each cluster

1. Corps sonore	2. Luminosité	3. Forme	4. Matière	5. Saveur	6. Caractère
Basson Clarinette Cuivré Coffre à musique Lyrique Nasillard Criard Résonant / Étouffé Explosif Percussif Martelé Fracassant Bruit blanc	Foncé Sombre / Clair Terne / Lumineux Mat / Brillant Opaque / Transparent Éclatant Cristallin	Aiguisé / Émoussé Coupant Tranchant Piquant Incisif Transperçant Florissant Bulbeux Rond Ovale Large Mince / Épais Plat Fermé / Ouvert Creux / Plein Carverneux Profond Entier	Cassant Vitreux, vitré Coulant Métallique Artificiel Dense Lourd Spongieux Ouateux Duveteux Doux Dur / Mou Sec / Mouillé Rugueux / Lisse Rêche Raboteux Fibreux Feutré Vaporeux Voilé Boisé	Chocolaté Crémeux Laiteux Mielleux Velouté Pulpeux Guimauve Collant Gras Chaud	Apaisant Chaleureux Confus Avec caractère Avec vitalité Vif Pétillant Fougueux Riant Pleurnicheur Nerveux Moqueur Rêveur
Bassoon Clarinet Brassy Music box Lyrical Nasal Shrill Resonating / Damped Explosive Percussive Martelé Shattering White noise	Dark Dark / Clear Drab / Luminous Mat / Bright Opaque / Transparent Shining Crystallin	Sharp / Blunt Slicing Pointed Incisive Transparent Blooming Bulbous Round Oval Large Thin / Thick Flat Closed / Open Hollow / Full Cavernous Deep Full	Brittle Glassy Flowing Metallic Artificial Dense Heavy Spongy Cottony Feathery Sweet Hard / Soft Dry / Wet Rough / Smooth Harsh Scraping Fibrous Felty Vaporous Veiled Woody	Chocolatey Creamy Milky Honeyed Velvety Pulpous Marshmellow Sticky Fat Warm	Appeasing Warm Muddled With character Vivacious Quick Sparkling Wild Laughing Whining Nervous Teasing Dreamy

Fig. 7.2 Classification of the adjectives into different sensory categories. Adjectives are given in French in the upper part of the table and in English in the lower part of the table.

regroups synonyms of a given adjective (from the compilation of several definitions in some cases). An adjective in bold is the centre of a cluster; the adjectives in the same cluster are its synonyms. Clusters delimited with a dashed line regroup lists of antonyms; the similarity between these descriptors is weaker than in the case of lists of synonyms.

On the map that finally resulted, the adjectives organized themselves along a main axis, from bottom left to top right on Fig. 7.3 (in French) and Fig. 7.4 (in English). This axis corresponds to plucking position. Hollow (*creux*) and dull (*mat*) sounds are found in the lower left-hand corner of the map; these sounds are obtained by plucking the string close to its middle. At the opposite extreme – in the upper right-hand corner – lie thin (*mince*) and nasal (*nasillard*); these sounds are usually obtained by plucking the string closer to the bridge.

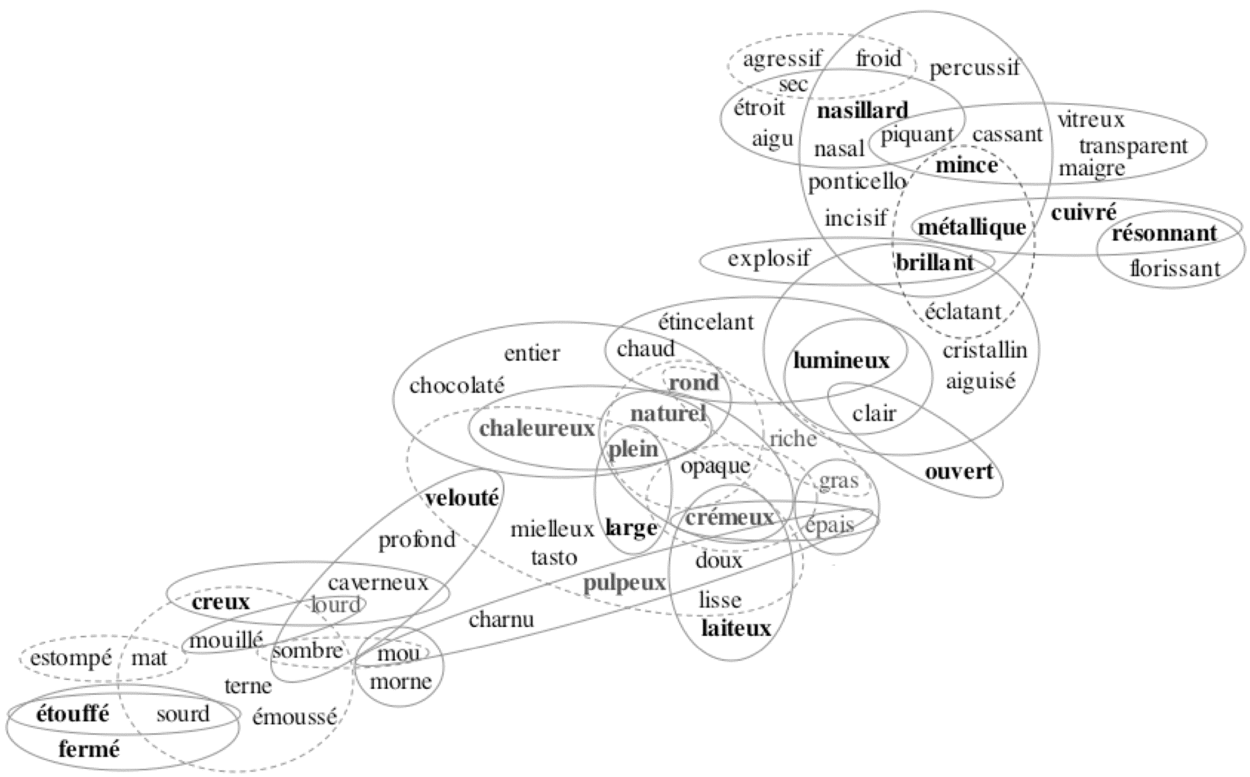


Fig. 7.3 Organization of timbre descriptors into clusters. Original French words are displayed. In the lower left-hand corner of the map, sounds are usually obtained by plucking the string closer to its middle. In the upper right-hand corner, sounds are usually obtained by plucking the string closer to the bridge. English translations are displayed on Fig. 7.4.

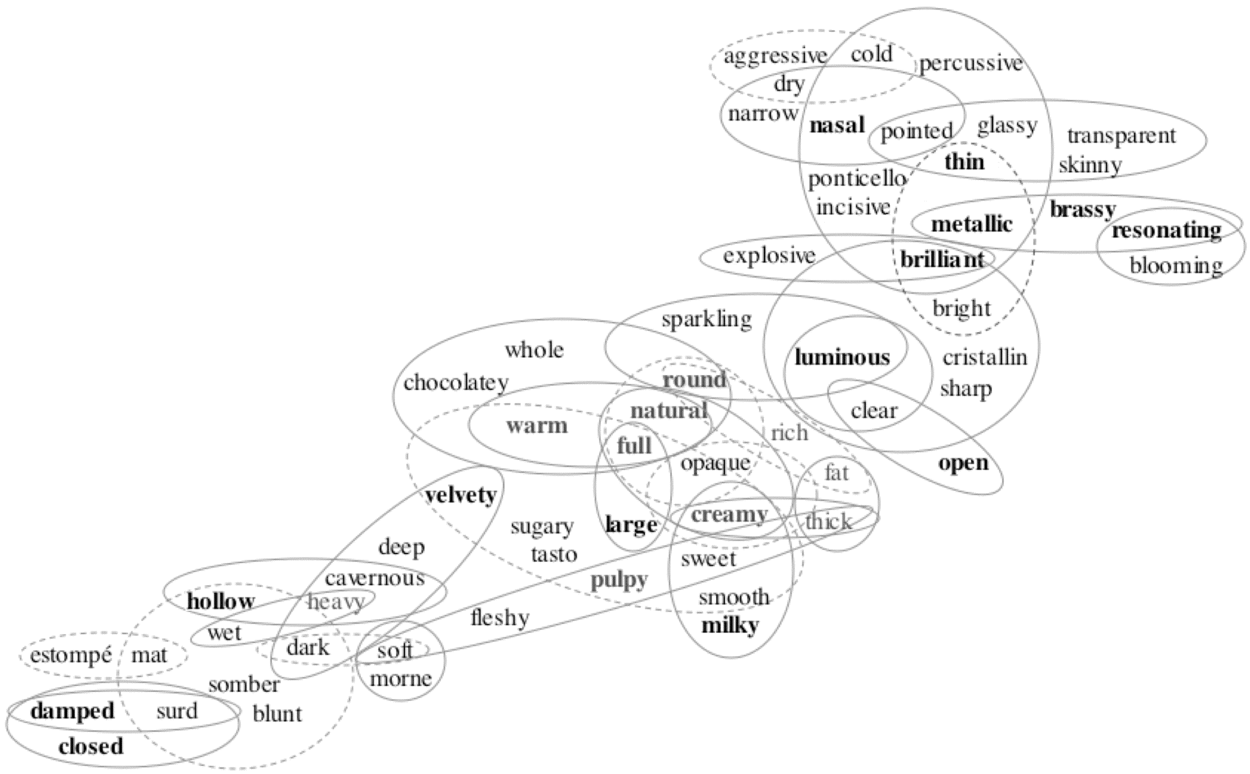


Fig. 7.4 Organization of timbre descriptors into clusters. English translations are displayed.

Some groups of synonyms did not connect with any of other adjectives, such as the cluster [*rugueux, raboteux, rêche*], which translate as rough, scraping and harsh respectively. These adjectives are not represented on the map.

7.2 Most common adjectives

For each of the 10 most commonly defined adjectives in the study, we compiled all the synonyms, antonyms, intuitive sound descriptions and associated gesture provided by the guitarists. The data is reported in this section presenting the adjectives from the brightest to the mellowest : dry, nasal, thin, metallic, bright, round, warm, thick, velvety and dark. The numbers in the left column (labelled with symbol #) of the tables below refer to the participant's identification number.

Dry

Synonyms : short, aggressive, percussive, firm, thin, mat, drab, raw, staccato

Antonyms : flowing, rich, legato, melodic

#	**Sound description**
4	Evokes a pizzicato.
6	A sound that does not travel. The noise of the string that just barely vibrates. It is more the plastic aspect of the string that is heard.
8	Small resonance. Sound with no sustain and without content. Gloomy as a dead tree. Obtained in the high register of strings 4, 5 and 6.
16	The sound is reduced to its attack but the pitch should remain defined.

#	**Gesture**
2	String attacked very close to the bridge with great strength. Resonance is shortened by damping the string with an other finger.
4, 13	Obtained by attenuating or damping the length of the note with the palm of the hand or by releasing the note.
8	String is attacked softly with last phalange souple.
16	Played with nails and very close to the bridge.

Nasal

Synonyms : thin, transparent, naily, dry, pointed, narrow, agressive, with a “twang”. “Nasal is very metallic” (# 18). “Nasal is thin and without depth” (# 6).

Antonyms : full, round, sober, natural/tasto, fat

#	**Sound description**
2	Sounds like someone is talking through nose. Not a pretty sound but humorous.
6	Contains a lot of high frequency harmonics.
9	Sounds like a duck or an oboe.
12	Sounds like the low and medium register of the harpsicord.

#	**Gesture**
6	Attacked with the nail only with the fingers and hand perpendicular to the string.
9	Played close to the bridge. Better illustrated on higher strings, particularly on the first string.
12	Slightly pulled upward and vigorously.
18	Plucked as close as possible to the bridge with the nail.

Thin

Synonyms : brittle, piercing but light, small, delicate, transparent, meagre, breakable, sharp

Antonyms : fat, heavy, large, round, full, harsh

#	**Sound description**
5	Characterizes the sound of a beginner player (because of the lack of control of the attack). Contains too many high frequencies. Sounds like the body of the guitar is not sollicited.
8	Sharp attack. Sound does not resonate much. Ressembles the sound of a banjo since it is a bit metallic. Sounds fragile as though it lacks assertion. Also sounds like an oboe.
12	Heard more in the high register.
17	Lacks richness compared to the natural sound of the guitar.
18	Timbre associated with the production of harmonics. Sounds like bells or celesta because of its short-lived resonance.
21	Sounds like the guitar has a very tiny body

#	**Gesture**
5	Played on the edge of the bridge with the finger perpendicular to the string.
8	Hand positionned between tonehole and bridge.
12	Attack angle towards the right, slightly pulled.
14	String is displaced upward during the attack with fingers perpendicular to the string.
17	Halfway between *metallic* and *bright* positions along the string.
21	Achieved with a pointed nail perpendicularly to the string.

Metallic

Synonyms : very nasal, very fat, very clear, very bright, silvery, brassy, brittle, percussive, harsh, ponticello, thin, pointed, sharp, rigid. "Too metallic is nasal" (# 5). "Extremely clear but agressif and nasal" (# 11).

Antonyms : mat, round, tasto, soft, velvety, warm, large, pulpy, creamy

#	**Sound description**
1	Sounds powerful
4	Sounds artificial and has a whistling vibration
6, 19	Evokes the harpsichord
8	Sounds a bit mechanical, like the sound of the hammer banging against an anvil
14	Evokes the banjo or the mandoline
15	Evokes the sound of steel drums

#	**Gesture**
1	Nails are almost perpendicular to the string, taking the shape of a hook.
1, 6, 8, 11, 13, 15, 21	Played very close to the bridge.
5, 18	Played closer to the bridge.
8	Played with the tip of the nail.
14, 17	Played as close as possible to the bridge (extreme ponticello).

Bright

Synonyms : clear, piercing, cristallin, pure, luminous, alive, sharp, firm. Clear AND round, brassy AND woody. "Very bright is metallic". "Bright is clear but not nasal nor weening".

Antonyms : mat, somber, hollow, muted, mellow, heavy, soft (as opposed to hard), dark, wet

Emotions : joy, distinction, rejoycing, solemn

#	**Sound description**
1	Very present and resonant and does not die rapidly.
3	Resonates without forcing.
4	Seems high-pitched even if the note is low.
6	Produces a lot of high harmonics.
8	The articulation of each note is heard distinctly.
12	Like a piercing bird sound that can be heard from far.
16	Has a good attack. This timbre is adequate for rapid passages.

#	**Gesture**
1	Played slightly towards the bridge.
3	Slightly diagonally with respect to the string above the edge of the tonehole towards the bridge.
4	Played close to the bridge with the tips of the nails.
6	Can be obtained only on the first 3 strings using the tirando technique.
8	Hand is placed above the edge of the tonehole closest to the bridge and string; attack is firm with last phalange rigid.
11	Played close (but not too close) to the bridge.
12	Pulled rapidly with strength and well articulated.
13	Played closer to the bridge with some good strength and articulation.
15	More easily obtained with new strings which are nervous and easily excitable (because more elastic).

Round

Synonyms : sweet, soft, mellow, creamy, rich, voluptuous, dense, heavy, velvety, thick, natural, lyrical, fat, warm, pulpy, “full but more metallic and veiled” (# 1).

Antonyms : narrow, thin, nasal, metallic, dry, bright

#	**Sound description**
4	Sound that rolls, very soft and limp.
6	Perfect balance between high and low harmonics, whether soft or strong.
8	Sounds like a bubble is coming out the guitar. Very soft sonority to the ear because it appears like a very homogeneous sonorous pillow. Like a cloud that gradually gets thicker and thicker.
9	Sounds like when the cork is removed form a champagne bottle but with a longer resonance.
11	Homogeneous and balanced sonority.
13	Heavy sound that tends to imitate a gong. Produces a bell-like modulation (as opposed to a wave-like modulation) projecting forward.
14	The sound of great guitarists.
19	Sounds like the attack is enveloping the resonance. Appropriate for slow and expressive melodies.

#	**Gesture**
1	With a slight apoyando.
6	Normal attack with as much nail as finger pulp.
8	Wrist turned towards the left. Fingers inclined inwards and hand above the middle of the tonehole.
9	Similar gesture than the one used to obtain a warm sound but with a longer preparation and a more firm last phalange.
13	A lot of pressure has to be applied to the strings with fingers slightly open. Slow attack.
14	String pulled down, close to tonehole with a 45 degree angle.
15	The nail should be cut in such a way that the distance between the initial contact point and the final falling point is maximized. The nail acts as a launching ramp. *Round* is halfway between *thin* and *thick* considering the sliding time of the string against the finger.
22	Played with the finger pulp only. Fingers slide upward and thumb slides downward against the string.

Comments : the analogy to the bottle neck sound (# 9) might refer to the labial resonator when forming a round vowel (see chapter 8).

Warm

Synonyms : round, chocolatey, mellow, velvety, sensual, natural. “round AND full”. “Perfect compromise between too nasal and too muted (étouffé)” (# 9).

Antonyms : angular, cold, glassy, dry, aggressive

Emotions or images sunset, joy, feeling of fullness

#	**Sound description**
1	Has a certain vitality and depth in the resonance.
6	Produces a lot of low harmonics but with still some energy in the high frequencies (characteristics of cedar guitar).
9	Evokes the comforting voice of a mother.
22	Emphasizes the medium-highs and the lows of the instrument.

#	**Gesture**
6	Played to the left with a lot of finger pulp and less nail. Plucking point is right above the middle of the tonehole.
8	Finger slightly inclined on string, string attacked softly above tonehole. More easily obtained with apoyando technique.
9	45^o angle between finger and string and slow vibrato.
22	Technique similar to the one for a round sound but even further away from the bridge.

Thick

Synonyms : dense, heavy, very fat, very round. “Warmest sound on a guitar” (# 9). “A thick sound is an exagerated round sound” (# 15).

Antonyms : transparent, thin

#	**Sound description**
14	Evokes the sound of the lute.
15	Evokes the steps of a giant or of an elephant.

#	**Gesture**
2	String attacked directly above the tonehole and on the left side of the nail.
8	Apoyando with wrist inclined towards the left.
14	Strings pulled downward, close to the tonehole, leaving the nail in contact with the string for the longest time (fingers almost parallel to the string).
15	Same as for a round sound but maximizing the gliding distance of the string on the width of the nail (small angle between finger and string). Plucked close to the tonehole.

Velvety

Synonyms : chocolatey, fudgy, feathery, fibrous, meliflous, sensual, rich, veiled, warm, very soft, very round. “Warm and round” (# 1). “Smooth and clear” (# 17).

Antonyms : bitter, crisp, dry, hard, rough

#	**Sound description**
18	A sound deprived of high frequencies.
22	Low volume and very soft attack.

#	**Gesture**
1	Played close to the nut (tasto) with “curbed” fingers.
4	Played with the pulp of the finger above the tonehole.
12	Apoyando, tasto on the second and third strings. Angle towards the left (like for a round sound).
13	Played towards the nut for a softer sound with the wrist turned slightly to the left without bending, softening the attack.
22	Same gesture as for a *warm* sound but the hand should be above the nut.

Dark

Synonyms : opaque, mat, carvernous, hollow, deep, velvety

Antonyms bright, luminous

#	**Sound description**
4	Brings out the base of one or several notes.
5	Refers to the heavy atmosphere of a composition.
6	Not aggressive. Possesses a lot of low harmonics and is rather soft.
9	Does not have much attack. Evokes the austerity and the seriousness of a church.
15	Sound that characterizes the guitars with a cedar top plate. It is like being at the edge of an immense bottomless pit.

#	**Gesture**
4	Obtained by playing with the finger pulp right above the frets and so that the highs do not resonate.
6	Played close to the nut (tasto) with a lot of finger pulp at the attack (less nail). So the attack should be very much on the left side of the nail in order to give the pulp as much expressivity as possible.
5	Played above the tonehole to get more roundness.
9	Needs a long preparation. The nail longly presses on the string and slowly slides in order to diminish the explosive effect of the attack.
15	Obtained on the lower strings.

Chapter 8

Phonetic Gestures Underlying Guitar Timbre Description

Voyelles

A noir, E blanc, I rouge, U vert, O bleu: voyelles,
Je dirai quelque jour vos naissances latentes:
A, noir corset velu des mouches éclatantes,
Qui bombinent autour des puanteurs cruelles,

Golfes d'ombre; E, candeurs des vapeurs et des tentes,
Lances de glaciers fiers, rois blancs, frissons d'ombelles;
I, pourpres, sang craché, rire des lèvres belles
Dans la colère ou les ivresses pénitentes;

U, cycles, vibrements divins de mères virides,
Paix des pâtis semés d'animaux, paix des rides
Que l'alchimie imprime aux grands fronts studieux;

O, suprême Clairon plein de strideurs étranges,
Silences traversés de Mondes et d'Anges:
- O l'Omega, rayon violet de Ses Yeux!

Arthur Rimbaud.

Contents

While investigating verbal timbre descriptors commonly used by guitarists (cf. Chapter 7), we discovered that some of them seem to refer to phonetic gestures: <u>open</u>, <u>oval</u>, <u>round</u>, <u>thin</u>, <u>closed</u>, <u>nasal</u>, <u>hollow</u>, etc. Indeed, as the magnitude spectrum of guitar tones is comb-filter shaped, we propose to consider the local maxima of this comb filter structure as acting like formants (cf. Chapter 5) and we compare them to the formants of vowels. For example, when guitarists describe a guitar sound as <u>round</u>, it would mean that it sounds like a vowel produced with a <u>round-shaped</u> mouth, such as the vowel [ɔ].

The next chapter reports on an experiment that was conducted in order to verify the perceptual analogies between guitar sounds and vocal sounds, based on the analogies that were found at the spectral level. In the experiment, participants were asked to associate a consonant to the attack and a vowel to the release of guitar tones. These analogies support the idea that some perceptual dimensions of the guitar timbre space can be borrowed from phonetics. This would imply that guitar sounds acoustically resemble voice sounds enough to engage a particular mode of timbre perception, what we call, the “phonetic mode” of timbre perception.

The present chapter is divided in four sections. First, the “phonetic mode” of timbre perception is introduced. In the second section, we compare the voice and the guitar from different points of view. The third section presents the acoustical characteristics of vocal

sounds. In the last section, we review how linguists and musicians describe the timbre of speech sounds. The interesting fact is that there exists a large set of common qualifying adjectives used for the description of guitar tones and speech sounds.

When referring to vowel sounds, we most often use the International Phonetic Alphabet. These symbols are placed between [ ]. We also use Slawson's sound colour notation (see Appendix B).

8.1 From speech perception to instrumental timbre perception

8.1.1 Articulation in speech

Articulation is primary in speech; it coexists with phonation that serves the secondary function of providing audibility (proof is that whispered speech is completely understandable). Articulation is superimposed upon the function of mastication.

When phonated, speech carries melodic information in the trajectory of the fundamental frequency of the vocal folds periodic excitation; when whispered, speech carries melodic information in the trajectory of formants.

According to Paget, language is a refinement of gesture: "In recognizing speech sounds, the human ear is not listening to music but to indications, due to resonance, of the position and gestures of the organs of articulation." [124] (p. 125). This theory goes along the lines of the motor theory of speech perception.

Paget makes the observation that a child concentrating intensely upon the mastery of a muscular act for a specific purpose will duplicate the act with other muscles. A child learning to write will "write with his tongue" at the same time; as he learns to tie his shoe laces, he will all but knot his tongue in a parallel process. If he so happens to make a sound during this process, either with vocalized or unvocalized breath, he will be speaking a word.

In order to command attention to his gesture, the primitive man would undoubtedly phonate loudly. Suppose he is making a gesture for "high" by raising his arm and unconsciously raises his tongue at the same time. He will say "AL". Coincidentally, "AL" does mean "high" in many languages or is found in words evoking height (names of mountains, for example) [124]. At first there will be gestures which have obvious concrete meaning like this one, and from these there will develop abstract and more symbolical meanings. With

phonation, the primitive man made his gesture audible, and it added emotional colour to it.

This theory places emphasis upon articulation. It is not making of sounds that is basic, but the movement of the speech organs. Accordingly, shifts in vowel formants would indicate movements more than they identify cavities. Even the different vowel colours are more articulation than phonation since they are a function of the resonators rather than the vibrators.

8.1.2 The motor theory of speech perception

The "motor theory" of speech perception was proposed by A. M. Liberman and his colleagues [133], [134]. In its most recent form, the model claims that "the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands that call for movements of the articulators through certain linguistically significant configurations" [134]. In other words, we perceive the articulatory gestures the speaker intends to make when producing an utterance [135]. A second claim of the motor theory is that there is an intimate and innate link between speech perception and speech production. Perception of the intended gestures occurs in a specialized speech mode whose main function is to automatically convert an acoustic signal into an articulatory gesture.

The proponents of this model have argued that it can account for a large body of phenomena characteristic of speech perception, including the variable relationship between acoustic patterns and perceived speech sounds, duplex perception, cue trading, evidence for a speech mode and audiovisual integration [134].

8.1.3 Non-speech mode vs speech mode of aural perception

When listening in a non-speech mode, the acoustic signals are received in the manner of musical sounds or natural noises; in the speech mode, acoustic signals are excluded from awareness, and only an abstract phonetic category is perceived [154].

Vowels and consonants have different linguistic and acoustic properties. The auditory parameters of speech (formant frequencies in steady-state vowels, for example) are analyzable by either hemisphere, whereas the linguistic features of the signal (those associated with consonants) can be extracted only by the hemisphere that is language-dominant. Nev-

ertheless, even relatively slow formant movements (as found in semi-vowels and liquids such as [l], [r], [w], [j]) may invoke the linguistic mode, as suggested by Haggard [130].

According to Reuven Tsur [153], humans are naturally tuned to the non-speech mode; as soon as the incoming stream of sounds reveals the slightest hint of linguistic information, however, we automatically switch to the speech mode: we digress our attention away from the acoustic signal to the combination of muscle movements that seems to have produced it, and from these elementary movements away to the phoneme sequence.

8.1.4 The poetic mode of speech perception

Tsur proposes a third type of speech perception: a poetic mode in which some part of the acoustic signal becomes accessible, however remotely, to consciousness. With Roman Jakobson's model of childhood acquisition of the phonological system [122], Tsur shows how the nonreferential babbling sounds made by infants form a basis for aesthetic valuation of language. He tests the intersubjective and intercultural validity of various spatial and tactile metaphors for certain sounds.

Tsur comments:

> In certain circumstances, in what we might perhaps call the *poetic mode*, some aspects of the formant structure of the acoustic signal may vaguely enter consciousness. As a result, people may have intuitions that certain vowel contrasts correspond to the brightness-darkness contrast, some other to the high-low contrast, or that certain consonants are harder than others. As a result, in turn, poets may more frequently use words that contain dark vowels, in lines referring to dark colors, mystic obscurity, or slow and heavy movement, or words depicting hatred and struggle. At the receiving end of the process, readers have vague intuitions that the sound patterns of these lines are somehow expressive of their atmosphere [153].

8.1.5 The phonetic mode of musical timbre perception

Similarly, we could propose a new type of instrumental timbre perception: a "phonetic mode" that consists in the unconscious perception of the combination of muscle movements of the speech organs that may have produced a similar instrumental sound. There might be sufficient linguistic information – such as the presence of formants in the magnitude spectrum – in the tones of a number of musical instruments that one may easily enter a sort of speech, or pre-speech mode, when listening to a performance.

8.1.6 Sound symbolism

None of the guitarists referred to vowel sounds in their written definitions (cf. Chapter 7); however, when asked in person to elaborate upon the definitions of round and thin, our collaborator guitarist Peter McCutcheon began to utter vowel sounds. Hence arose the idea to look for formants in the comb-filter-shaped magnitude response of guitar tones (as defined in Chapter 5). The vowels that illustrated a round sound were produced with a round-shaped mouth (e.g. the open [ɔ] as in the word *lock*). A thin sound – which is often regarded as an antonym for round – was vividly depicted with vowels obtained by spreading sideways the lips (e.g. a [i] sound as in the word *tea*); to produce this vowel, the mouth has a thin shape. In the same stream of thought, the guitarist described an oval sound with [Λ], as in the name R*u*ssel.

This connection between guitar sounds and vowels does not appear to be within the realm of consciousness among classical guitarists who, as a species, generally favour a more metaphorical and abstract vocabulary to describe timbres (as was the case with all written timbre descriptor definitions collected and reported in Chapter 7). Before drawing explicit parallels between guitar sounds and vowels, here are some additional facts that support this reasoning.

Some classical guitar masters will occasionally sing phrases with vocables (nonsense syllables) in an effort to communicate a timbre to their students. The guitarist Manuel Barrueco is known for asking his disciples to “make their guitars sing” without further explanation as to what he has in mind. Bass guitar players often vocalize the bass line when communicating with one another. An even more explicit connection between speech sounds and instrumental sounds is found in the realm of percussion instruments, especially with the North Indian Tabla tradition, an oral tradition which uses a system of vocables to name drum sounds. In a recent study, Patel and Iversen [151] tested the hypothesis that the vocables are a case of sound symbolism (onomatopoeia). Analysis revealed that acoustic properties of drum sounds were reflected by a variety of phonetic components of vocables. More generally, drum performers seem to use onomatopoeia to refer to different types of strokes.

The brightness (or spectral pitch) of vowels has unconsciously been a part of human knowledge for a long time. This is attested by the onomatopoeic words in our language. For example, we regard *moan, groan, shout, yell, scream, shriek* as ranging from low to

high in spectral pitch [147].

8.1.7 The vocal quality of formant glides

Musicians continuously aim to emphasize the vocal quality of their instruments. This quality is more effectively obtained with "varying formants" rather than with "steady formants". In fact, the characteristics of speech sounds go beyond the presence of formants in vocal sounds to the capability of articulating smoothly between vocal sounds, resulting in a "formant glide". This change in the spectral envelope occurs when saying "wah" by smoothly switching from [u] to [a]. The peaks of the spectral envelope shift, and the quality of the sound changes without any pitch change in the tone.

The most familiar example of formant glide is the "wah-wah effect" created by trumpet players and electric guitarists. Schneider explains how a formant glide effect can be achieved on an acoustic guitar: "If one of the lower three strings is plucked and then lightly damped, the spectral envelope is transformed into one very closely resembling the *oo* spectral envelope. The damping finger absorbs all the harmonics except the fundamental and the first few overtones. This, the reverse of the *wah* effect, can be vocally described as *ah*-oo" [30].

8.1.8 Correlating musical timbres with vowels

In his thesis [158] which investigates related analytical techniques for music and language, Youngblood reports that some pitches played on the bassoon sound like a spoken [æ]. He suggests that "a full-scale attempt to correlate musical timbres with vowels could be undertaken, and the results would be useful to composers, music educators, and speech and hearing specialists. It is possible that a person who has difficulty distinguishing between two musical instruments might also have difficulty distinguishing between the vowels with which these instruments correlate".

8.2 Comparing voice and guitar acoustical systems

8.2.1 The "singing" guitar

The ideal timbre that guitarists set to achieve is a <u>warm</u> and <u>round</u> sound. The purpose of the interaction with the instrument is to impregnate the tone with as much of an organic

quality as possible so that it alludes to a sound produced by the human voice. In order to achieve this quality, guitarists have to compensate for the rigidity of the instrument acoustics. The walls of vocal resonators are made of soft tissues (tongue, cheeks, palate, etc.), while resonators of most musical instruments are made with hard materials. In the voice acoustical system, there is a weak coupling between excitation and resonator. Some resonating cavities are variable in shape and size. In the guitar acoustical system, string and body are two resonators strongly coupled through the bridge. The body is fixed in size and shape. Despite all of these differences, the performer seeks ways to achieve timbral manipulations that occur in speech.

8.2.2 Vowel-like resonances in musical instruments

In his book *Sound Color*, Slawson reviews research that investigated "vowel-like resonances in some musical instruments" [154]. He notes that "most musical instruments have sources that are driven by the resonance systems of the horns, strings, or membranes that make up this instrument" and that "[t]here is little in those systems, apparently, that is vowel-like" [154] (p. 157). He refers to Jansson [8] who compares the bow-string system to the vocal source and the resonance box to the vocal-tract filter system; this analogy is not very successful. Slawson rules out musical instruments as models for sound colour because of the strong coupling between source and filter: "[t]here may be some basis for studying the sound colour of musical instruments if other decoupled resonance systems are discovered" [154] (p. 157).

Nonetheless, we believe that in order to establish perceptual analogies between vowel sounds and guitar sounds, it is not necessary to find strong similarities between the structures of the acoustical systems (i.e. between the causes of the sounds). It might be sufficient to find similarities between the acoustical signatures of the sounds produced by these systems (i.e. between their effects), regardless of their cause.

In a source/filter modelling of the vocal production system, the source consists of the vocal folds that generate the glottal excitation, and the filter corresponds to the cascade of resonators formed by the vocal tract (oral, labial and nasal cavities). Vowels are recognized based on the frequency location of the formant regions.

As illustrated on Fig. 8.1, the guitar may also be decomposed into a source (the strings) coupled to a filter (the body) via a bridge. In an attempt to draw parallels between

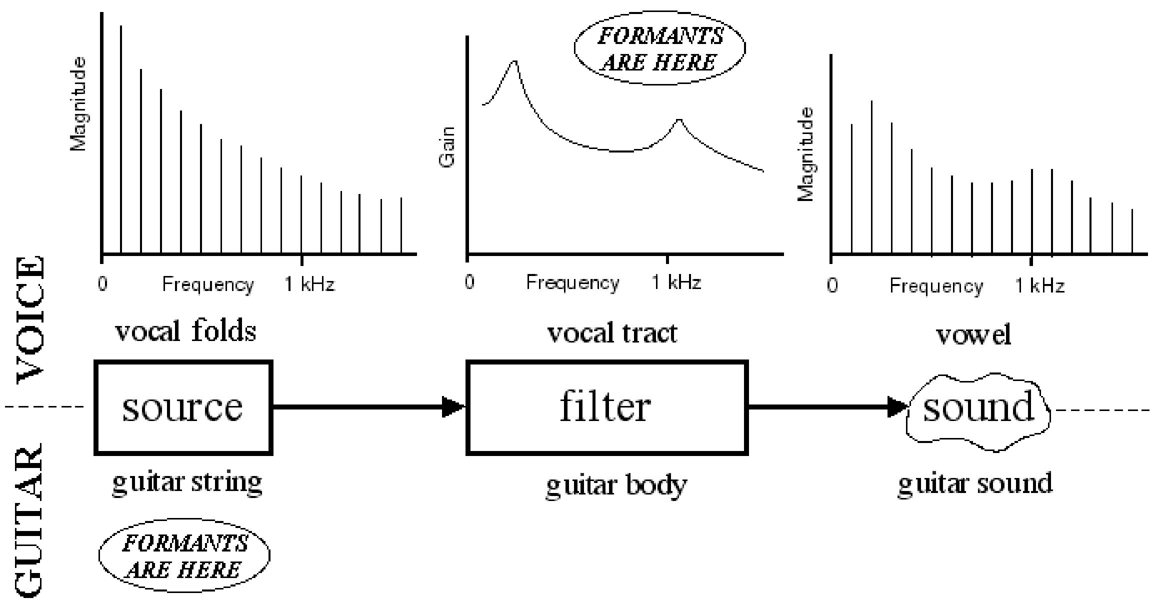


Fig. 8.1 In the source/filter model of the vocal mechanism, the formant structure lies in the resonator. In the source/filter model of a guitar, the formant structure lies mainly in the source.

the guitar and the voice acoustical systems, one would expect that the formant structure belongs to the filter in both cases. We suggest that the vowel-like formant structure is actually due to the localized plucking excitation point along the string (resulting in a comb filter effect) rather than to the main resonances of the body which occur at quite low frequencies, around 100 or 200 Hz.

As shown in Chapter 5, the comb filter effect, inherent to any localized excitation point along the string, is characterized by odd-numbered formant frequencies ($F_2 = 3 \times F_1$, $F_3 = 5 \times F_1$, etc.). Some vowels show similar patterns in their magnitude spectrum since the vocal tract is, in first approximation, a tube closed at one end, that also favours odd-numbered resonant frequencies. This situation occurs for the neutral vowel, as illustrated on Fig. 8.2.

To summarize, vowels and guitar tones often display similar acoustic signatures, although the systems that produce them are structurally different (the latter is a coupled system whereas the former is not). In order to establish perceptual analogies between vowel sounds and guitar sounds, we believe it is sufficient to find similarities between the acoustical signatures of the sounds, regardless of their cause.

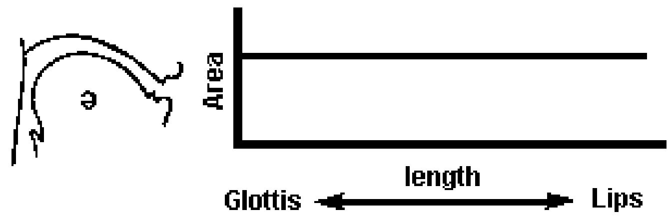


Fig. 8.2 Neutral vowel represented by stylized vocal tract configuration and area functions [123].

8.2.3 Vocal strings and sounding board

The voice is not a stringed instrument, although certain parallels can be drawn. For instance, the vocal "cords" change their pitch by changing their tension, as do strings. But strings do not alter their tension while they are being played. In fact, the vocal folds function much more like the lips of a trumpet player.

Other analogies between the voice and a stringed instrument that is sometimes used by singers are even less justified, namely references to the hard palate and other bony surfaces as "sounding boards". The voice has no strings, or if one considers vocal cords as such, there is no bridge to any part of the body that might be called a sounding board. Furthermore, the palate is too small to act as a sounding board, and it is muffled in a soft, fleshy covering [147].

8.3 The voice acoustical characteristics

Some timbre descriptors for the classical guitar refer to vocal characteristics, such as roundness and nasality. This is why we present in this section the acoustical basis of these two features of the vocal production system.

8.3.1 The singing voice

The singing voice is a wind instrument. The actuator is the wind supply in the lungs of the singer. The vibrator is in the larynx, or voice-box. The excitation is produced by the vocal folds, behaving as a valve periodically interrupting the flow of air coming from the lungs The resonators are the laryngeal, oral, labial and nasal cavities.

8.3.2 Formant structure of vowel sounds

The recognizable quality of the vowel sounds is due to the existence of formant regions, which are frequency ranges where the sound is enhanced by the cavity resonances of the human vocal tract. The spectral envelopes corresponding to three different English vowels are displayed on Fig. 8.3.

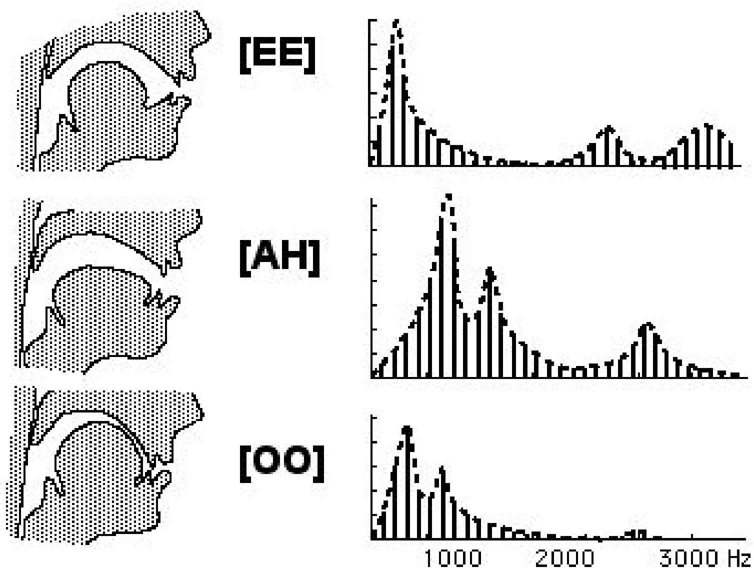


Fig. 8.3 Spectral envelopes corresponding to three different vowels.

A vowel's timbre depends on the following elements:

- the number of active resonators (among the laryngeal, oral, labial and nasal cavities);
- the shape of the oral cavity (determined by the general position of the tongue in the mouth – front, central or back positions);
- the size of the oral cavity (depending on the degree of aperture of the mouth).

8.3.3 The mouth as a resonator

The mouth is also called the oral or buccal cavity. The boundary between the oral and the pharyngeal cavities is marked by the solf palate at the top, the pillars at the sides, and the tongue at the bottom.

The position of the tongue determines whether the mouth and the throat will function as one large air chamber or as two resonators, and how the cavities will vary in size. The size is also influenced by the position of the jaw, and the shape and dimensions of the orifice are a function of the lips and teeth [147]. The parts of the mouth are also used in the production of consonants.

According to Delattre, the lowest frequency formant is tuned by the overall opening of the bucco-pharyngeal resonator. When the tongue and jaw are raised, as for [i] and [u], the low formant is at its lowest point; and as the jaw drops and the mouth opens, the low formant rises toward [ɑ]. The second formant is tuned by the "lengthening" of the mouth cavity and is lowered as a result of tongue backing and lip rounding [128] .

8.3.4 The lips as a resonator : the roundness in the voice

If the lips are rounded and pushed forward, a third labial resonator is formed; if, on the other hand, the lips are spread sideways or pressed against the teeth, no labial resonator is formed. The presence of this resonant frequency in the spectrum may therefore be correlated with the perceptual attribute of roundness.

8.3.5 The nose as a resonator : the velvet in the voice

The nasal cavity itself is not adjustable, so the control consists entirely of shunting it in or out of the resonance system. If the soft palate is raised, air does not enter the nasal cavity and passes mostly through the oral cavity; a vowel produced in such a way is an oral vowel. If the soft palate is lowered, air can pass through nose and mouth simultaneously, producing a nasal vowel. The perception of a sound as nasalized depends on the ratio of the size of the opening into the nasal cavity and the opening into the oral cavity. When the nasal port is large relative to the oral port, then nasality is perceived.

The tone as resonated by the nose is a honky, muffled sound. In classical singing, the closure of the naso-pharynx is usually complete. Vennard mentions that "A small seasoning of nasality is sometimes desirable to give the voice a velvety quality" [147]. Also, nasality is the characteristics of certain consonants, represented by the symbols [m], [n], and [ŋ].

The nasal tract has its own resonant frequencies or formants, the nasal formants, which vary from speaker to speaker due to the large variation in side and shape of nasal cavity.

One consequence on the magnitude spectrum of the vowels appears to be at low frequencies, in the vicinity of the first formant [131]. There is usually an upward shift in the first formant frequency due to the presence of a low frequency antiresonance just below F_1, which tends to make the peak in the spectral envelope in the region of F_1 appear between 50 and 100 Hz higher than it would normally be.

Antiresonances occur whenever there is a side branch in the main acoustic pathway, as it is the case when the soft palate is lowered, allowing the air to pass through both oral and nasal cavities. The most general effect of adding nasal resonance to oral resonance is an overall loss of power. The antiresonances decrease energy at specific frequencies, thereby reducing and sometimes eliminating some low intensity formants from the acoustic signal. The general attenuation of the signal is also reflected in the broadening of all the formant bandwidths and flattening of spectral peaks. The association between broad bandwidths and nasal sounds was noted by Jakobson & al. [121] (p.39). Also, the degree of nasalization heard depends on the amount of acoustical impedance in the oral and nasal cavities. As a result, high vowels (such as [i]) are generally perceived as more nasal than low vowels (such as [ɑ]).

The fact that the magnitude spectrum of a guitar tones displays broad peaks might explain why guitar tones are generally perceived as nasal.

8.3.6 The larynx as a resonator : the brilliance in the voice

Among singers, brilliance refers to the *ring* in the voice. The *ring* of the voice is the presence of strong overtones averaging around 2800-2900 Hz for men, and about 3200 Hz for women [147]. This is the third formant also called the singer's formant because it is much more present while singing than while speaking [146]. In fact, the singer's formant comes from the reunion of several formants (3d, 4th and 5th), one of which might be associated with the resonance of the chamber of the lower larynx.

This *ring* has various characteristics that associate it with the larynx. According to Vennard, although two formants are sufficient to identify vowels and while some vowels (especially [i]) are more ringing than others, "the presence in strength of the *ring* marks the fine singer" [147] (p. 129).

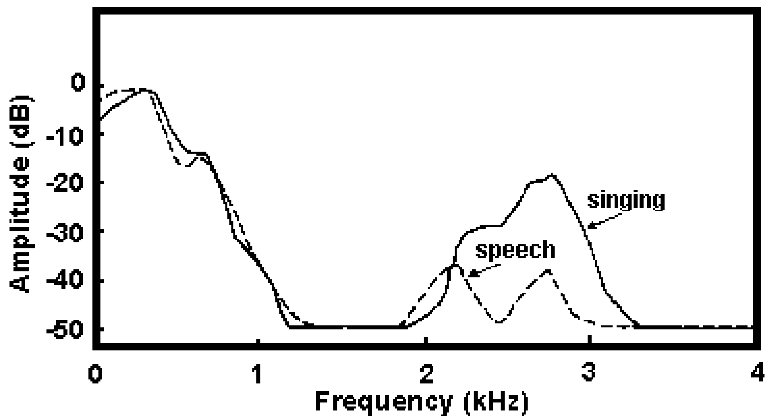


Fig. 8.4 Spectrum envelopes of the vowel [u] spoken (dashed curve) and sung (solid curve) by a professional opera baritone singer. The amplitudes of the harmonics between 2 and 3 kHz, approximately, give a marked envelope peak in singing. This peak is called the singer's formant, typically occurring in all voiced sounds produced by Western classical male singers and altos. (Adapted from [145]) [146].

8.3.7 The texture of the resonators

The hardness or softness of the surfaces of a resonator encourages or discourages high overtones. The softer the walls of the resonator, the greater the attenuation of high frequencies and the larger the bandwidth of resonances. Therefore, soft walls, absorbing the rapid, short-wave vibrations, make a tone sweeter, mellower. Hard walls make the tone more brilliant by reflecting the high partials. Vocal resonators have various surfaces. Most of them are fleshy and thereby soft, but the hard palate has a bony structure near enough to the surface to make a difference. Vennard points out that, just before the tone emerges, it may pass through either a soft, fleshy orifice – as in the vowel [u] – or a sharp, hard-edged orifice – as in the vowel [i], especially when it is "smiled" due to the hardness and sharpness of the teeth [147]. Vennard adds that "if the throat and the root of the tongue are stiff when singing with force, the tone which emerges is ringing to the point of being <u>metallic</u>" [147] (p. 155).

An other aspect to consider is the warm and humid air resonating in the vocal tract, increasing visco-thermal losses and therefore widening formant bandwidths.

8.3.8 Coupling between source and filter in the voice acoustical system

In the voice acoustical system, the excitation and the resonator are therefore not completely decoupled. The vowels have specific effects on laryngeal function through the extrinsic musculature. The glottal rattle requires a loose glottis, and is much more difficult to perform on either [i] or [u] than it is while the resonators are forming an [ɑ]. The ventricules of the larynx are larger for [i] than [ɑ] with the same loudness, especially when the vowels are whispered. The vowel [ɑ] requires much less pressure to produce, and the [u] slightly less than the [i] [143]. Singers are trained to give a little more energy to the vowel [i], partly because the mouth is likely to be more closed, and also because this vowel contains more energy in the high frequencies [147]; nevertheless, the produced global power is usually weaker for [i].

8.4 The description of speech sounds

An interesting fact is that there exists a large set of qualifying adjectives used for the description of both guitar tones and speech sounds. This section reviews how linguists and musicians describe the timbre of speech sounds; a particular attention is payed to vowels.

8.4.1 Physiological description of speech sounds

The principal physiological factors that are considered when distinguishing vowels from one another are [120]:

- Movement of the tongue forward or backward with the jaw held steady. Example: [æ - a - ɔ] as in *panned - pond - pawned*.
- Movement of the mouth and jaw from almost closed to fully open with the tongue held steady. Example: [i - e - æ] as in *mean - mane - man*.
- Rounding or non-rounding of the lips with the tongue and jaw held steady. Example: [ü - i] as in German *Tür-Tier*.
- Opening or closing the passage to the nasal cavity with the tongue and jaw held steady. Example: [ɔ̃ - o] in French *bon - beau*.

Consonants differ according to the following principal criteria:

- Presence or absence of voicing (vocal folds vibration). Example: *din* - *tin*.
- Complete or partial obstruction of the air flow. Example: *tin* - *sin*.
- Closure or non-closure of the velum. Example: *dip* - *nip*.
- Surmounting or circumventing the obstruction. Example: *rip* - *lip*.

In addition to these clearly perceived differences, trained phoneticians can hear relatively subtle sound differences, like the differences between the *p*'s of *p*in, s*p*in, and na*p*kin.

8.4.2 Distinctive features of speech sounds

The distinctive feature theory was proposed by Jakobson, Fant and Halle in 1951 [121] and then later revised and refined by Chomsky and Halle in 1968 [119]. The theory codifies certain long-standing observations of phoneticians by hypothesizing that many sounds of speech can be categorized based on the presence or absence of certain distinctive features: whether the mouth is open, whether there is a narrowing of the vocal tract at a particular place, whether a consonant is aspirated. Those properties are the features that characterize and distinguish the phonetic content of a language. The theory can be applied, with only slight modifications, to all human languages throughout the world. Jakobson, Fant and Halle detected twelve inherent distinctive features in the languages of the world. They present those features as binary oppositions (d.f. stands for "distinctive feature"):

- Fundamental source features
  - Vocalic vs non-vocalic [d.f. 1]
  - Consonantal vs non-consonantal [d.f. 2]
- Secondary consonantal source features
  - Envelope feature
    - Interrupted vs continuant [d.f. 3]
    - Checked vs unchecked [d.f. 4]
  - Strident vs mellow [d.f. 5]
  - Supplementary source: voiced vs voiceless [d.f. 6]

- Resonance features
  - Compact vs diffuse [d.f. 7]
  - Tonality features [d.f. 8]
    - Grave vs acute [d.f. 9]
    - Flat vs plain [d.f. 10]
    - Sharp vs plain [d.f. 11]
  - Tense vs lax [d.f. 12]
  - Supplementary resonator: nasal vs oral [d.f. 13]

8.4.3 Slawson sound color

In his book *Sound Color* [154], Slawson addresses the following question: "How can one aspect or dimension of sound color be held constant as other dimensions of sound color are varied?" He answers by first designating three of the distinctive features of vowels (compactness [d.f. 7], acuteness [d.f. 9] and laxness [d.f. 12]) as candidates from which to derive dimensions of sound colour. Then he determines equal-value contours for the distinctive features as shown on the next figure. To hold sound colour constant with respect to one dimension, Slawson recommends changing the values of the first two formant central frequencies F_1 and F_2 in such a way as to remain on one of the equal-value contours of this dimension.

- OPENNESS (replacing the term COMPACTNESS given in [121]) is named for the tube shape with which it is correlated. The approximate acoustic correlate of OPENNESS is the frequency of the first resonance.
- ACUTENESS reflects its connotation of high or bright sound. It increases with increasing frequency of the second resonance.
- LAXNESS is said to correspond to a relatively relaxed state of the articulatory musculature. The equal LAXNESS contours are closed curves on the (F_1, F_2) plane centered on the maximally lax point. This central point correspond to the formant values that would arise, in theory, from the vocal mechanism in the position to which it is automatically brought just before beginning to speak [119]. This is the neutral position

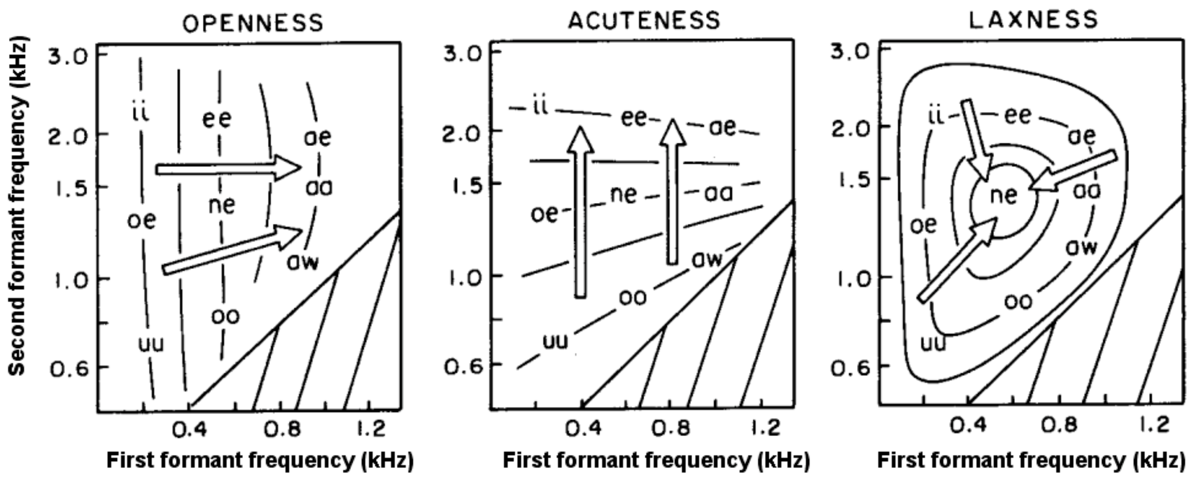


Fig. 8.5 Contours of equal OPENNESS, equal ACUTENESS and equal LAXNESS.

of the vocal tract which can be best approximated by a single tube closed at one end. Since a tube of length L closed at one end can only resonate at frequencies for which L is an odd multiple of one quarter wavelength and since the average length of the vocal tract of males is about 17.5 cm, the resonances appear at approximately 500, 1500, 2500 Hz, etc. A tense vowel displays a greater deviation from the neutral formant pattern (Supplement – Tenseness and laxness by R. Jakobson and M. Halle in [121].)

Slawson notes that "if we can identify certain primitive features of speech that serve some pre-speech function, we have reason to consider their inclusion among the features of sound in general and of sound color in particular" [154] (p. 61). In some sense, Slawson claims that the dimensions of OPENNESS, ACUTENESS and LAXNESS are fundamental biological features that are part of the auditory processing of all sounds. The colour of a sound is determined by its value on each of the dimensions, and its phonetic category in the speech mode may be determined by which side of a critical point on the dimensions its sound color lies.

Though Slawson intuitively recognized that sound colour distinctive features can be applied to the auditory processing of all the sounds, and therefore to the sounds produced by musical instruments, he did not propose any specific applications since he maintained that clear analogies between vocal sounds and instrumental sounds were not possible.

8.4.4 Metallic quality of some consonants

Phonetically, while [p] and [t] are "diffuse" consonants, [k] is characterised as "compact", or more abrupt. The consonant [k] is often associated with a metallic or brassy sound. As noted by Gaver [86], the sounds made by vibrating wood decay quickly, with low frequencies lasting longer than high ones, whereas the sounds made by vibrating metal decay slowly, with high-frequency lasting longer than low ones. In addition, metal sounds have partials with well-defined frequency peaks, whereas wooden sound partials are smeared over frequency axis. Tsur [153] adds that the opposition "well-defined frequency peaks" - "smeared over frequency axis" may be perceived as corresponding to the compact-diffuse opposition in the traditional phonetics domain, characterising the [k] - [p, t] opposition. Tsur remarks that there is nothing metallic in the velum, the place of articulation of the [k]. It is the acoustic features of [k] that render it more metallic than [p] or [t]. "This can explain why we hear a clock tick-tocking rather than, e.g., tip-topping. [k] is better suited than [p] or [t] to imitate the metallic click of a clock."

In this explanation of the metallic quality of [k], the frequency contents of the sounds are compared. Since plosive consonants are transients, it could be more appropriate to compare the rise times. In fact, what metallic sounds and a guttural plosive such as [k] have in common is a very short attack duration. It can be argued that this is a better explanation of the association between the sound [k] and the metallic quality.

8.4.5 Description of vowels by singers

Singers devote much effort to the control of vowel quality. The vocabulary they use to qualify vowel timbre is really close to the vocabulary used to qualify instrumental timbre (e.g. round, pointed, dark, open, ...). Our primary source for this section is *Singing. The Mechanism and the Technic* by William Vennard [147].

Round vs pointed

Vowel [ɑ] (as in "c*a*lm") is part of the *round vowels* family and [i] (as in "b*ee*t") is part of the *pointed vowels* family. The qualities associated with the [i, e, ɑ] series are bright, cool, forward, pointed, high. The opposite qualities are associated with the [u, o, ɑ] series: dark, warm, back, round, deep. The vowel [ɑ] belongs to both groups.

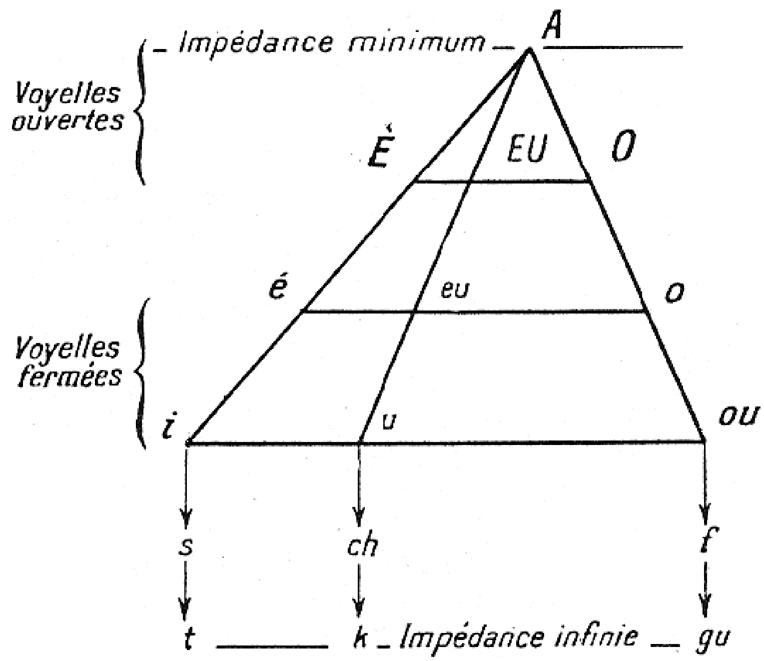


Fig. 8.6 Vocal triangle by Hellwag (1781) [142].

Vennard explains that "singing the *round* vowels tends to lower the voice box, and therefore they may sound better and actually rise to higher pitches than the *pointed* vowels" [147] (p. 154).

The open and round vowel [ɑ]

Vennard points out that the vowel [ɑ] (as in "c*a*lm") predominates in the voice literature. It is the most fully resonant sound in language and it shows the greatest variety of possible colour. He adds that in [ɑ], "brightness and mellowness are equally balanced" [147] (p. 145). To produce [ɑ], the pharynx is distended comfortably, the jaw is dropped, the tongue is low and grooved. "[ɑ] may be considered an [uh] (neutral vowel) that has been beautified by proper resonation" [147] (p. 131).

It is interesting to note that the plucking region of predilection for guitar sounds is where round and open sounds are produced. A great number of adjectives qualify guitar sounds of this type (as shown on Fig. 7.3 from the previous chapter).

The pointed vowel [i]

The vowel [i] (as in "b*ee*t") sounds brilliant regardless of how it is produced. The frequency of the forward cavity of the vowel [i] has a frequency near that of the collar of the larynx, which is said to produce a resonance around 2800 Hz. Incidentally, this is the resonant frequency of the chamber of the outer ear, which means that this formant need not be very loud to sound loud, since the ear is very sensitive around that frequency.

In [i], the *ring* is higher pitched and may overpower the lower partial. This is what happens when the [i] sounds strident, white or nasal. Vennard explains that in this event, the tongue may be too stiff, or the teeth may comprise too great a part of the aperture, either of which conditions favour high partials at the expense of low ones [147] (p. 145)

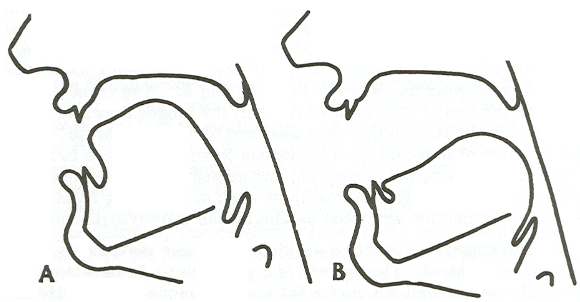


Fig. 8.7 Contrasting tongue positions : generalized outlines based upon X-rays by G. O. Russell [126]. **A** resembles various X-rays to which are applied the words: barbaric, flat, metallic, piercing, pinched, tight, voix blanche. **B** resembles various X-rays of professional singers and others, to which are applied the words: forward placement, mellow, resonant, soft. [147].

The pear-shaped [ɑ]

This analogy, which seems to have originated with Lilli Lehman [144], summarizes the idea that a good vowel will have "forward placement" while filling the entire throat. Vennard describes the analogy in the following terms [147] (p. 149) :

> "The stem of the pear is the teeth; it is the *focus* of the tone, which is a means of getting desirable quality. The stem itself is undesirable, and used only in vocalizing. It is nasal and twangy in its most extreme form, but the whole fruit grows from it. The small part of the pear is in the mouth and is the

> brilliance of the tone. The pear swells into something large and mellow, that can be felt throughout the entire pharynx and is limited only by the singer's ability to enlarge this organ"

This description fits the vowel [ɑ] most aptly. The placement of the other vowels is described with differently shaped objects : the [o] has less of the mouth resonance, and is more "like an orange", the [e] is more "cone-shaped" than a pear, and the [i] is even further pointed.

White vs dark voices and vowels

Singers call dark, dark brown, grey or muddy a timbre that lacks high partials. The French expression, voix blanche ("white voice") designates a singer's voice which contains high partials at the expense of low ones [147]. The resonance dichotomy is forward brilliance as opposed to mellow depth. Also in singing, roundness is often associated with depth.

Garcia states that the voice has two timbres, clair and sombre (p. 5). As to the physiology of these timbres, Garcia elaborates that for clear or open timbres, the larynx is high and the soft palate low whereas for dark or covered timbre the larynx is low, and the velum high, and the pharynx vigorously rounded. When *timbre clair* is exaggerated, the voice becomes white, shrill, and screeching (*blanche, criade, glapissante*). When *timbre sombre* is exaggerated, the voice is covered, choked, muffled (*couverte, étouffée, sourde*) [147] (p. 121).

This white/dark opposition characterizes the different tradition of singing pedagogy. The Italian tradition of pedagogy is to emphasize "forward placement", which makes for great brilliance and flexibility; whereas the Germanic teachers have been more likely to emphasize a deeper production, the "stroke of the glottis", which makes for fuller tone and more power, such as Wagnerian opera demands.

Pursed lips lower the pitch of the resonators by decreasing the diameter of the orifice, giving it soft edges, and adding a "neck" to the resonator. According to Vennard, this positioning of the lips is necessary to produce the dark vowels, [o] and [u] [147] (p. 119).

If the lips are drawn back, as in an exaggerated smile, the edge of the orifice becomes the teeth, which draws out high partials. If the opening is made larger, the pitch of the resonator raises. The acoustic effect is the exact opposite of darkening, and is called whiteness, or voix blanche. The hard edges of the teeth are sometimes employed to give

brilliance to an otherwise breathy tone, as for example, a falsetto.

The origin of the adjectives round, open, closed, etc. is quite obvious, but the opposite pair dark and clear are not explained as easily. Perhaps the vocal tract is felt as a dark cave. If the sound seems to emerge from a point very low and far inside the vocal tract, the sound is perceived as dark. If the sound seems to emerge from a higher point, closer to the opening of the mouth, the sound is perceived as clear and white. The acoustic shadowing is perceived as an optical shadowing.

Emotional connotations of the vowels

There are various theories of the origin of language. One theory involves the concept that the vowels are instinctive expressions of emotion, from which other, more specifically communicative expressions have evolved [147].

It can be said generally that the high formant vowels are more elated, whereas low formants are more sombre. Vennard explains that if a voice student swallows his tones, the teacher will often suggest to sing them more gaily, more happily. If the student sings too whitely, the teacher would suggest a more sober, more profound sound. A sad song might be sung as if the [e]'s were [ø] and [i]'s were [y]. In an exultant song, [o] becomes [ɔ], etc. [147] (p. 147).

Chapter 9

Listening to Guitar Sounds as Vocal Sounds

Between
What I think
What I want to say
What I think I say
What I say
What you want to hear
What you think you hear
What you hear
What you want to understand
What you think you understand
What you understand
There are ten chances that we will have trouble communicating.
But let us try anyway...

Bernard Werber
*(*from *L'encyclopédie du savoir relatif et absolu).*

Contents

This chapter reports an experiment that was conducted in order to verify the perceptual analogies between guitar sounds and vocal sounds, based on the analogies that were found at the spectral level. In the experiment, participants were asked to associate a consonant to the attack and a vowel to the release of guitar tones. These analogies support the idea that some perceptual dimensions of the guitar timbre space could be borrowed from phonetics.

9.1 Application of the distinctive features of speech to guitar sounds

9.1.1 Guitar sound subspace in a vowel space

The trajectories that we plotted with a dotted line on top of Slawson's equal-value contours for distinctive features of speech in Fig. 9.1 correspond to the relationship $F_2 = 3\,F_1$, which

is found for the first two local maxima of a comb filter frequency response[1]. The first formant frequency is calculated with:

$$F_1 = \frac{l f_0}{2p} \tag{9.1}$$

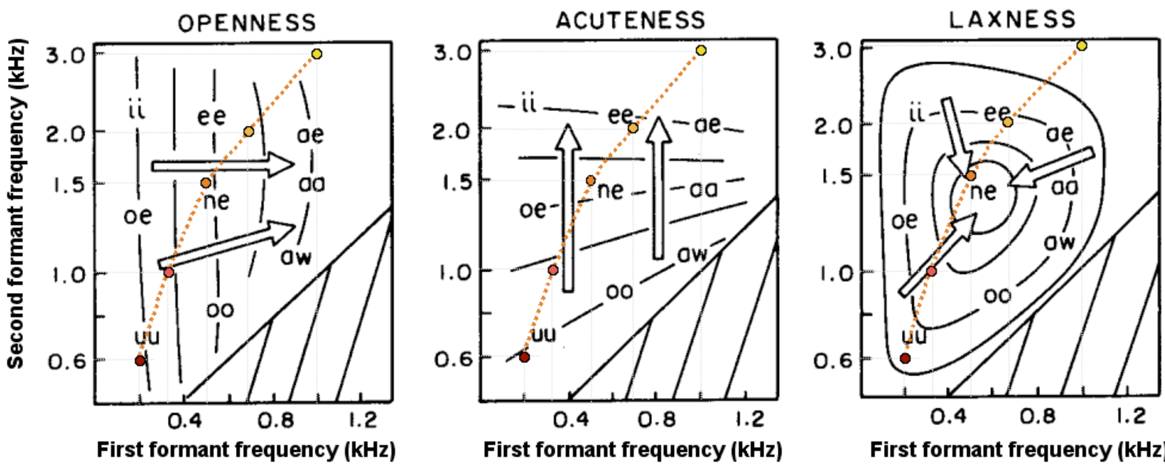


Fig. 9.1 Equal-value contours for three distinctive features of speech in the (F_1, F_2) plane [154] with superimposed *guitar vowels* trajectory (dotted line) corresponding to the relationship $F_2 = 3\,F_1$.

In that way, one can see which "vowels" are obtained by varying the plucking position from the middle of the string (**uu** region) to the bridge (**ee** region or further up depending on fundamental frequency of the string). For a given string, the absolute plucking position p will determine the vowel colour, regardless of the note that is played. In fact, the formant frequencies F_n are constant for a given absolute plucking position p on a given string, regardless of the note that is played because the product $l f_0$ in Eq. (9.1) is a constant for a given string and equals half the speed of sound c along the string. Therefore, F_1 can also be expressed as the ratio $c/4p$. As a result, vowel colour is maintained for any note on a given string, except for relative plucking position $R = 1/2$ which is the case of an odd-harmonic only spectrum, perceived as a distinct timbre.

The table below gives the first formant frequency calculated with Eq. (9.1) for the six strings of a guitar tuned with the standard tuning, together with the closest sound colour

[1] The curve $F_2 = 3\,F_1$ is not a straight line because the F_2 axis is not linear.

and corresponding IPA symbol, for a string length $l = 60$ cm and a plucking position $p =$ 12 cm from the bridge (normal plucking position).

String	Note name	Tuning frequency	First formant frequency for $p = 12$ cm and $l = 60$ cm	Closest sound colour	IPA symbol
6	E (Mi$_2$)	83 Hz	$(30 \times 83)/12 = 207.5$ Hz	**uu**	u (b*oo*t)
5	A (La$_2$)	110 Hz	$(30 \times 110)/12 = 275$ Hz	**oe**	ø (b*ö*se)
4	D (Ré$_3$)	146 Hz	$(30 \times 146)/12 = 365$ Hz	**oo**	o (b*oa*t)
3	G (Sol$_3$)	202 Hz	$(30 \times 202)/12 = 505$ Hz	**ne**	ɜ (th*e*)
2	B (Si$_3$)	248 Hz	$(30 \times 248)/12 = 602$ Hz	**ee**	e (b*ai*t)
1	E (Mi$_4$)	330 Hz	$(30 \times 330)/12 = 825$ Hz	**ae**	æ (b*a*t)

Table 9.1 For each string, the frequency of the first comb filter formant is calculated for a pluck 12 cm from the bridge ($F_1 = lf_0/2p$).The closest sound colour (based on Fig. 9.1) is also given.

We see that different vowels correspond to different strings. Darker vowels (such as [u]) will be heard on lower strings. Clearer and thinner vowels will be heard on higher strings.

The G-string (lowest of the three nylon strings) is particular. In the normal plucking position, it produces a neutral vowel, also called the dull vowel by singers. This was confirmed by our collaborator Peter McCutcheon who complained about the G-string as always "too dull". Another characteristic of this string is that the whole guitar-vowel trajectory is covered when plucking position is varied from the middle of the string to the bridge. At its midpoint, the string (with $f_0 = 202$ Hz) starts with a **uu** ($F_1 = 230$ Hz, $F_2 = 700$ Hz) sound since $F_1 = f_0 = 202$ Hz and $F_2 = 606$ Hz. Plucking a bit closer to the bridge (for example at 27 cm from the bridge on a 60 cm string), F_1 gets even closer to the first formant central frequency of **uu**.

$$F_{1_{(\mathrm{p}\,=\,27\ \mathrm{cm})}} = \frac{lf_0}{2p} = \frac{60 \times 202}{2 \times 27} = 224 \text{ Hz}$$

Very close to the bridge (for example at 5 cm), the plucked G-string produces a thinner and more nasal tone, close to a [ɛ̃].

$$F_{1_{(\mathrm{p}\,=\,5\ \mathrm{cm})}} = \frac{lf_0}{2p} = \frac{60 \times 202}{2 \times 5} = 1212 \text{ Hz}$$

It is not possible to produce dark vowel such as **uu** on the first string (with $f_0 = 330$ Hz).

The darkest vowel on the first string is obtained at its midpoint, where $F_1 = f_0 = 330$ Hz. Guitarists confirm that the first string is often "too thin".

9.1.2 Phonetic gestures underlying guitar timbre description

On Fig. 9.2, we drew "mouth shapes" associated with the different degrees of OPENNESS and ACUTENESS. A closed and acute vowel (such as **ii**) is represented with a thin horizontal ellipse. The neutral vowel **ne** (in the center) is moderately open and moderately acute.

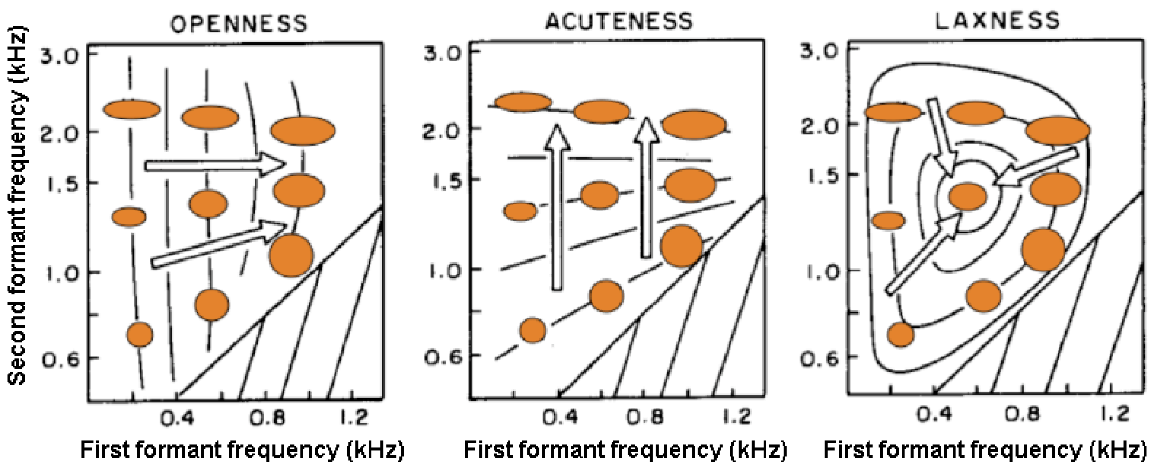


Fig. 9.2 Mouth shapes associated with vowel colours centred on the corresponding (F_1, F_2) points.

We propose to apply the three distinctive features of speech to guitar sounds in order to explain the origin of some of the adjectives that guitarists use to describe timbre. For example, the adjectives closed, round, large, open could indicate different degrees of OPENNESS. The adjectives thin and round would be opposites along the ACUTENESS dimension. A warm or chocolatey sound would likely be associated with the maximally LAX point. In fact, a warm sound likely evokes the sound that one makes while exhaling warm air, usually with the vocal tract in a neutral position. Finally, a hollow or cavernous sound would actually sound like the [u] vowel produced as the mouth forms a hollow cavity.

9.1.3 Holding sound colour constant

Since each string corresponds to a distinct vowel colour for a given absolute plucking position, we could ask by how much should the plucking position change from a string to an

adjacent string in order to maintain the sound colour constant.

In the standard tuning, all strings are a perfect fourth (5 semitones) apart except between the 3$^{\text{d}}$ string and the 2$^{\text{d}}$ string where there is only an interval of a major third (4 semitones). Knowing that the first formant frequency F_1 is proportional to the fundamental frequency f_0 and inversely proportional to the absolute plucking position p, it can be concluded that the absolute plucking position has to increase or decrease proportionally to f_0. Let α be the transposition ratio from one string to the next, say 4/3 for a fourth. The fundamental frequency f_0 is multiplied by α when switching to the higher-frequency string and p should be multiplied by the same factor α in order to keep the first formant frequency constant. In fact, it is easy to verify that

$$F_1 = \frac{l\alpha f_0}{2\alpha p} = \frac{lf_0}{2p}$$

To summarize, if two adjacent strings are a fourth apart,

$$\Delta p = \frac{4p}{3} - p = \frac{p}{3}$$

and if the two adjacent strings are a third apart,

$$\Delta p = \frac{5p}{4} - p = \frac{p}{4}$$

Example: a tone is plucked 15 cm from the bridge on the second string. When switching to the first string (a fourth higher), the pluck has to be $15 \times 4/3 = 20$ cm away from the bridge to keep F_1 constant ($\Delta p = 5$ cm).

Guitarists do not usually compensate for the change of timbre. It would be physically quite difficult and unpractical since the right hand would have to bend to the right so that the index finger could pluck the highest string closer to the bridge. Most guitarists play "to the left" with the hand in the axis of the forearm. As a result, higher frequency strings are plucked slightly closer to the bridge than lower frequency strings (2-3 cm difference between index and ring finger). Thus, this playing technique accentuates the timbre differences between strings instead of attenuating them.

There is one situation in which guitarists try to compensate for the change of timbre: when playing a rapid scale apoyando, the hand drifts away from the bridge while switching

from lower strings to higher strings.

9.1.4 Relationship between formants

Peterson and Barney have found different formants for the same vowel in men, women, and children. It appears that the vowel is a Gestalt in which the frequencies of the formants occupy a large role, but in which relationships between the formants also contribute to recognizability. Hence, the illusion of a certain vowel sound may be possible with unusual formant frequencies [136].

9.1.5 Timbral continuity from note to note on the same string

Schneider notes : “When a melody is performed on a single string and the right hand stays at the same plucking position, the spectrum of each note is different. This is best illustrated by playing an octave diatonic scale on a single string with the right hand at one-sixth of the length of the open string throughout the scale. When the octave is reached, at the 12th fret, the right hand will be plucking at one-third of the vibrating length, having plucked at a different percentage of the length for each note of the scale.” [30] p. 37.

It is true that the spectrum changes from note to note. The relative magnitude of harmonics can be dramatically different. Granted, at the beginning of the scale example heretofore mentioned, the 6th harmonic is attenuated (as well as all its integer multiples) and at the end of the scale, an octave higher, the 3d harmonic is attenuated. However, it is still the same absolute frequency since the frequency has been doubled when reaching the octave.

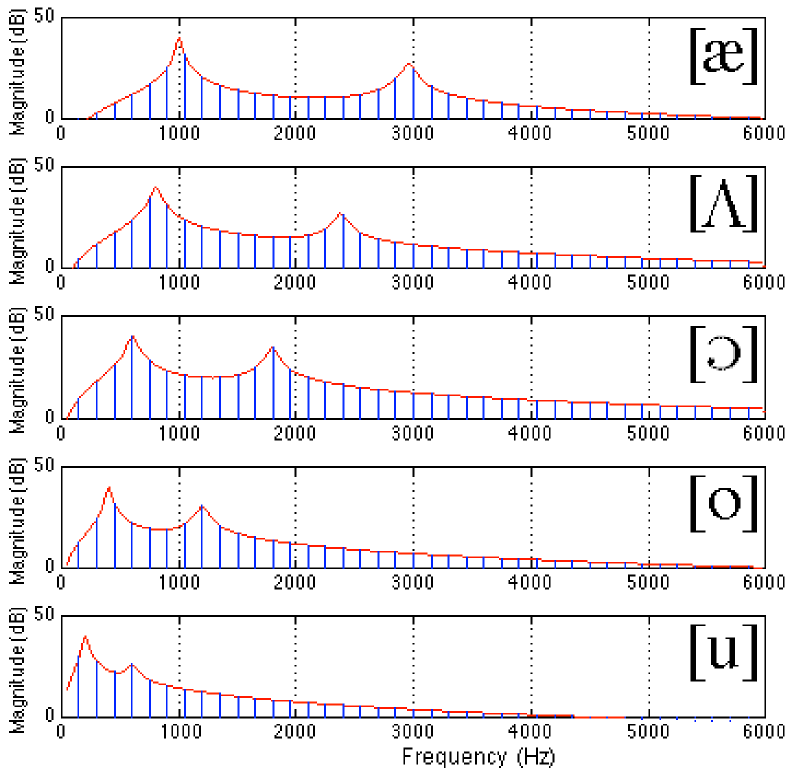
As shown in Chapter 5, a fixed absolute plucking position induces a constant absolute location of the maxima and minima in the magnitude spectrum. It is as though the spectral envelope were fixed, while the harmonics move around under it. This is illustrated by the Fig. 5.3 and Fig. 5.4 in Chapter 5. Note that this behaviour emulates the spectral behaviour of the voice.

9.2 Associating non-sense syllables to guitar tones

When guitarists are asked to associate vowel sounds to guitar tones obtained with various plucking positions ranging from near the bridge to closer to the midpoint of the string,

they agree on the sequence [ɛ̃], [ɛ], [ɑ], [ɔ], [o], [u] (when imposed to choose among these simple French vowels) from the bridge to the middle of the string.

We have synthesized vowel-like sounds characterized by formants located at the frequency locations of the maxima of a comb-filter structure: the second formant frequencies equals three times the first formant frequencies ($F_2 = 3 \times F_1$ with $F_1 = 1000, 800, 600, 400$ and 200 Hz) as shown on Fig. 9.3.



	f	A	BW
F_1	1000	0	40
F_2	3000	-13	90
F_1	800	0	70
F_2	2400	-14	80
F_1	600	0	70
F_2	1800	-6	80
F_1	400	0	40
F_2	1200	-10	80
F_1	200	0	40
F_2	600	-18	60

Fig. 9.3 Spectral envelopes with two formants of "guitar vowels". The table provides the central frequencies **f** (in Hz), the amplitudes **A** (in dB) and the bandwidths **BW** (in Hz) of the two formants F_1 and F_2.

This sequence of sounds simulates the narrowing of the comb-filter structure when moving the plucking position from the bridge to the midpoint of the string. The synthesized sounds are perceived as close to [æ] (as in "b*a*t"), [Λ] (as in "b*u*t"), [ɔ] (as in "b*ou*ght"), [o] (as in "b*oa*t"), [u] (as in "b*oo*t"). At this point, attention can be given to the shape of the mouth forming these vowels. When plucked close to the bridge, the string produces

a sound that is associated with a thin-shaped mouth. Moving closer to the tonehole, the mouth seems to open up to a round shape. Then, from the tonehole to the midpoint of the string, the mouth closes progressively while maintaining a more or less round shape. At midpoint, the guitar sound lacks all even harmonics. In fact, perceptually, the sound is generally described as hollow and some guitarists qualify it as a bassoon sound. The guitarist Alexandre Lagoya calls it a "son tuyau" [pipe sound] [26].

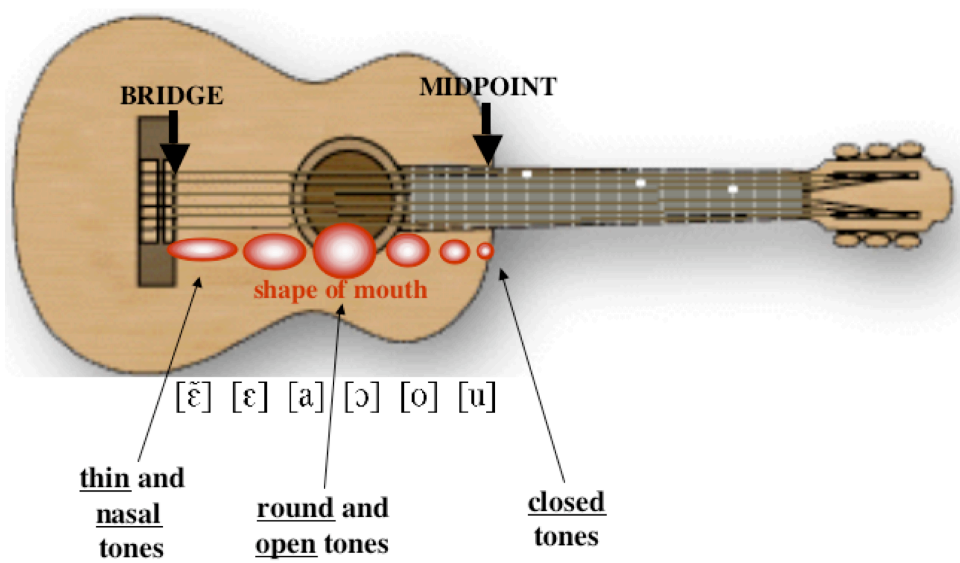


Fig. 9.4 Phonetic gestures associated with timbres with different plucking positions (the guitar was drawn by Matti Karjalainen – used with permission).

Note that the transitions from a thin-shaped mouth to a round-shaped mouth and then to a closed mouth are the same transitions one continuously goes through when imitating the sweeping flanging effect of a landing airplane, for example.

In order to confirm whether the vowel analogies could be perceived by non-guitarists, we conducted the following experiment.

9.2.1 Experiment

Nine French-speaking non-professional musicians and non-guitarist performers were asked to sing nonsense syllables that they deemed perceptually close to guitar tones, associating a consonant to the attack and a vowel to the release of the tone. To produce the stimuli, a professional guitarist was asked to play the same melody with different timbres. We selected

four variations of the performance which were described by the guitarist as ponticello, brassy, round and tasto.

The main instrumental gesture parameter that was varied to obtain the different timbres was the plucking position (from very close to the bridge to over the finger board). The angle of the plucking finger also differed, positioned closer to a perpendicular to the strings for brighter timbres. This correlation between the two gesture parameters was necessary to preserve the naturalness of the plucking techniques. The ponticello timbre was played 5 cm from the bridge with fingers perpendicular to the strings (thus a 90^o angle between the fingers and the string). The brassy timbre was obtained by plucking the string 8 cm from the bridge with a 60^o angle. The round timbre was obtained by plucking the string 13 cm (close to the tone-hole) from the bridge with a 45^o angle. The tasto timbre was obtained by plucking the string 20 cm for the bridge with a 30^o angle.

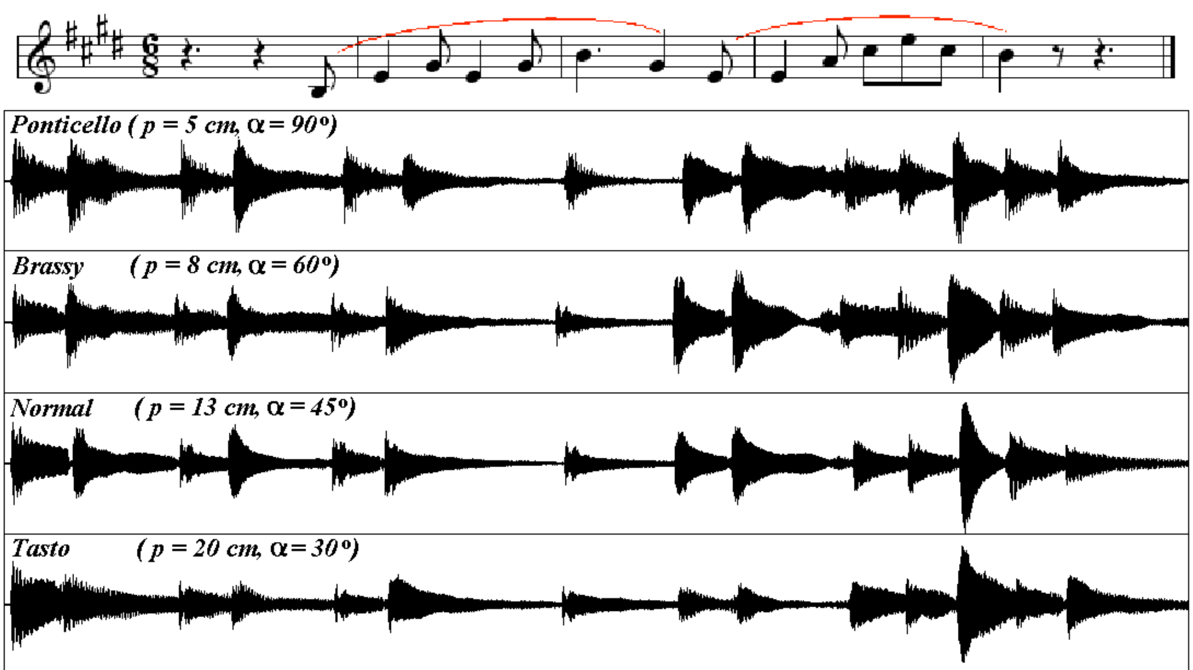


Fig. 9.5 Time-domain representation of the 14 tones of a melody played with 4 different timbres. The melody is an excerpt from the piece *L'encouragement* for two guitars by Fernando Sor (1778-1839).

The participants were not disclosed any information about the way in which the tones were produced nor about the timbres that were intended by the guitarist. The excerpts were

	1	2	3	4	5	6	7
Note	Si_3	Mi_4	$Sol\#_4$	Mi_4	$Sol\#_4$	Si_4	$Sol\#_4$
String	3 (Sol_3)	2 (Si_3)	1 (Mi_4)	2 (Si_3)	1 (Mi_4)	1 (Mi_4)	1 (Mi_4)
Fret	4	5	4	5	4	7	4
	8	9	10	11	12	13	14
Note	Mi_4	Mi_4	La_4	$Do\#_5$	Mi_5	$Do\#_5$	Si_4
String	3 (Sol_3)	3 (Sol_3)	2 (Si_3)	1 (Mi_4)	1 (Mi_4)	1 (Mi_4)	2 (Si_3)
Fret	9	9	10	9	12	9	12

Table 9.2 Fingering for the 14 notes of the melody. The name of the note is given, together with the string on which the note is played (number and name) and the associated fret number.

presented in a random order. The participants were free to replay the excerpts themselves in any order they pleased and as often as they needed. Additional information was collected by means of free verbalization of the participants.

9.2.2 Results

Table 9.3 reports the syllables provided by the nine participants for the different timbres. Plosive consonants ([k], [g], [t], [d], [p], [b]) were associated with the attack portion of the guitar sounds, while nasal or oral vowels were associated with the release portion of the guitar tones. Some participants provided two syllables because they found that the

	Ponticello	Brassy	Round	Tasto
Participant # 1	tɛ̃	tœ̃	ta	tø
Participant # 2	tɛ̃-ti	d[ɛ̃-ɑ̃]	bɑ̃	bwɔ̃
Participant # 3	kɛ̃	pɑ̃	dɔ	bɑ̃
Participant # 4	kɛ̃	tɛ-tɔ̃	tɔ	dɑ̃
Participant # 5	[k-t]ai	[d-p]aw	dɑ-dɔ	dɑ̃
Participant # 6	kɛ̃	gœ̃	tɔ	dø
Participant # 7	dɛ̃-kɛ̃	t[ɑ̃-ɔ̃]	dɔ̃-tɔ̃	gu-du
Participant # 8	kɛ̃	tsɑ̃-pɑ̃	dɔ-tɔ	θɔ̃
Participant # 9	kɛ̃	tɛ	ta	bu

Table 9.3 Non-sense syllables chosen by the 9 participants for the 4 timbres.

timbres differed from note to note (example : [tɛ̃-ti]). The other participants were able to determine a single syllable that would be most representative of the whole melody. In some cases, the consonant was hard to define and was said to be "*between a* [k] *and a* [t]", for example. They are notated between square brackets in the table (e.g. [k-t]). Similarly, intermediate vowels were provided, such as a nasal vowel between [ɛ̃] and [ɑ̃], notated [ɛ̃-ɑ̃] in the table[2].

- For the ponticello timbre (p = 5 cm from the bridge), most participants noted a strong nasal quality of the tones and chose the French nasal vowel [ɛ̃]. The consonant [k] seemed to evoke the metallic quality of the attack.
- For the brassy timbre (p = 8 cm from the bridge), the tone was described nasal but not as nasal as for the *ponticello* timbre. The chosen vowel was also more open (French nasal vowels [ɑ̃] or [œ̃], or English diphthong [aw]). The consonant [t] was chosen most often for the attack. Therefore, as the plucking point moves away from the bridge, the vowel becomes less nasal and opens up.
- For the round timbre (p = 13 cm from the bridge), the tones seemed to be perceived as oral and rounder vowels ([ɑ], [ɔ]). The consonant [d] was most often associated with the softer attack of the tones.
- For the tasto timbre (p = 20 cm from the bridge), all participants noted the hollow and closed quality of the tones, referring to the vowel [u]. While making the vowel sounds with their mouth, some participants mentioned that they felt they had to create a large space inside their mouth and close the lips. For the attack, softer consonants [b] and [d] were often chosen as well as the English consonant [θ], evoking the sound of the friction of the finger against the string.

The results of this experiment support the analogies that were found at the spectrum level. Considering the harmonic portion of the guitar tones (i.e. the decay), the tones are perceived more ACUTE and more NASAL when plucked very close to the bridge. At the other extreme, closer to the middle of the string, the tones are perceived more CLOSED. In the normal position, the tones are perceived ROUND and OPEN.

[2] The IPA symbols for the French nasal vowels are [ɛ̃] as in 'v*in*', [œ̃] as in 'br*un*', [ɑ̃] as in 'bl*an*c' and [ɔ̃] as in 'b*on*'.

With regards to the attack portion of the guitar tones, the further away from the bridge the string is plucked, the softer the attack is perceived. From harder to softer, the unvoiced consonants are [k - t - p] and the voiced consonants are [g - d - b - θ].

9.2.3 Voiced legato and unvoiced staccato

In the case of an unvoiced (surd) plosive consonant, the vocal tone is broken and a noise is inserted; in the case of a voiced (sonant) plosive consonant, there is no interruption in the vocal folds' periodic excitation, especially when singing. For example, singing [pa-pa-ti-pa-pa-ta] sounds less legato then [ba-ba-di-ba-ba-da]. Applying this principle to a melodic line played on the guitar, a given attack might be perceived "unvoiced" in a staccato passage and "voiced" in a legato passage. In order to verify this, we asked our collaborator guitarist Peter McCutcheon to vocalize a guitar line staccato and then legato. As expected, he used [t] in the first case (singing [ta-ta-ti-ta-ta]) and [d] is the second (singing [da-da-di-da-da]). I pointed out to him that he had changed the consonant. He was convinced he did not. He repeated the exercise in a more attentive state of mind and realized his different uses of the two consonants.

Scripture [127] reports a similar story. In studying some records by the tenor Caruso, he found that the singer frequently kept his vocal folds vibrating during sounds like [t] and [k]. This was done unconsciously; Scripture relates that Caruso was incredulous and indignant when the peculiarity was pointed out to him, yet the general effect of his singing was smoother on account of the peculiarity. Scripture suggests that it is often not only easier but also more pleasant to voice consonants between vowels: "The expression 'aha' with a voiced 'h' is the milder and more agreeable word; 'aha' with the unvoiced 'h' is an expression with more vigour, aggressiveness, and unpleasantness" [127] (p. 7).

Vennard agrees with the function of the noises in the vocal tone. He considers that legato singing is perceived as mild and agreeable. To add vigour, aggressiveness or unpleasantness to singing, the speech noises should be exaggerated into noticeable interruptions [147].

In the results of the experiment reported in Table 9.3, we noted that voiced plosive consonants were consistently chosen for the tasto timbre (the only exception is participant # 1 who focused his attention only on vowels and chose [t] for the four timbres). In fact, for the tasto timbre, the melody was played particularly legato.

From this we learned that when studying instrumental timbre perception, it is not suffi-

cient to investigate the characteristics of individual tones. Articulation plays an important role in the way a timbre is perceived and interpreted. It is interesting to note that a similar observation has long been made in the domain of speech processing.

Chapter 10

Comparing Music and Language Elementary Units

De la musique avant toute chose,
Et pour cela préfère l'Impair
Plus vague et plus soluble dans l'air,
Sans rien en lui qui pèse ou qui pose.

Il faut aussi que tu n'ailles point
Choisir tes mots sans quelque méprise :
Rien de plus cher que la chanson grise
Où l'Indécis au Précis se joint.

C'est des beaux yeux derrière des voiles,
C'est le grand jour tremblant de midi,
C'est, par un ciel d'automne attiédi,
Le bleu fouillis des claires étoiles !

Car nous voulons la Nuance encor,
Pas la Couleur, rien que la nuance !
Oh ! la nuance seule fiance
Le rêve au rêve et la flûte au cor !

[...]

Paul Verlaine.
(from *Art Poétique*)

Contents

Motivated by the analogies found between vocal sounds and musical instrumental tones, we reconsider, in this chapter, the comparison between phoneme units and scale units as expressed by different authors and then shift to a more acoustically founded comparison between phonemes – the sounds of a language – and what we call *sonemes* – the sounds of a musical performance.

10.1 Phoneme vs note

10.1.1 Comparison based on functional value

Phonemes are a set of universally accepted and understood symbols used to describe the sounds of a language as it is spoken. Phonemes transcribe the timbral features of a language, but not the pitch, the dynamics, the duration nor the speed of articulation.

The notes on a score indicate the pitch and duration of the sounds the performer must play. Scores generally include dynamics as well. In Western instrumental music, timbral features are rarely notated.

Springer [155] states that, in both phoneme systems and scale systems, "each [...] constituent [member] derives its functional value from its relationship to all other members.". From there, Youngblood suggests that the equivalent of a phoneme unit in music would be a scale unit. He draws the parallel between pitch classes and phonemic classes. This analogy is based on the fact that phonemes and pitches both are discrete units of their respective systems, and both have relative functional value. This is a very limited and traditional Western view of language, in which the prosodic information is attributed a small importance; and of music, in which timbre is repeatedly dismissed as a secondary parameter.

From an acoustical point of view, a phoneme unit ought to correspond to an aspect of timbre rather than to an aspect of melody. Different vowels can be produced with a given pitch as instrumental timbre can be varied with a given pitch. The pitch contour of a melody finds its speech counterpart in the form of an intonation contour.

10.1.2 Phonemes and notes as they are heard

Other parallels between phonemes and notes have been established. Nattiez [150] describes the phoneme as a "discretized" unit of language and the note as a "discretized" unit of music. They are "discretized" rather than discrete units, since a phoneme removed from its context has little meaning on its own, just as a note has little meaning when removed from its piece of music [148]. Here, Nattiez considers the note "as it is heard" and not as it is notated on a score (i.e. reduced to its pitch and duration values).

Wishart's view is that "the melodic stream is pitch-disjunct and may be articulated by timbral colouration. [And that the] language stream is timbre-disjunct and may be

articulated by pitch inflections" [156].

These two approaches deal with the sonic relationships between the phoneme, as the "discretized" unit of language, and the note, as the "discretized" unit of music [148]. This implies a continuum of articulated sounds or utterances in which language and music exist as they are heard [148]. This continuum is illustrated on Fig. 10.1.

Levman indirectly supports this idea in his discussion on the origins of music and language where he states that the differences between the performance of music and language are of degree, not of kind. Pitch, dynamics, duration and speed of articulation are all used in speech and in music, but their gamut is wider in music [149] (pp. 151-152). Music may have evolved out of language and songs would then be exaggerated speech. It is also possible that music and language developed from the same 'proto-faculty', and that as language became more expressive of ideas rather than of feelings, accent decreased as consonantal articulation increased [149] (pp. 147-149).

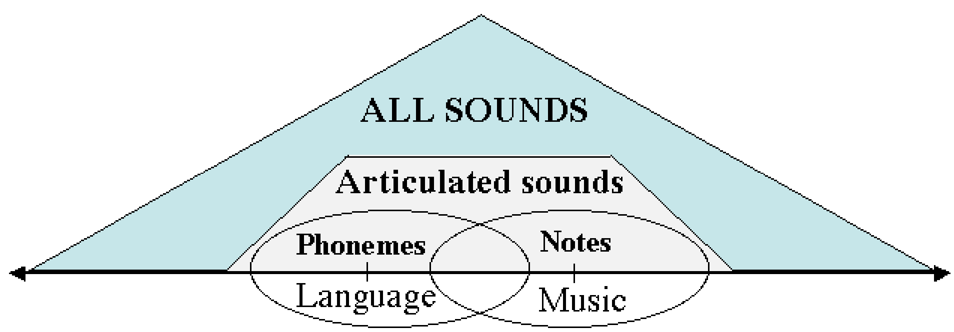


Fig. 10.1 Elementary units of language and music in the continuum of all sounds [148].

We would like to propose a refinement of this parallel between music and language elementary units. This refinement is instigated by the following observation: the phoneme symbol only transmits timbral information and the note symbol only indicates pitch and duration. If music is considered as it is heard, the term"note" leads to confusion. To remediate this, we propose the term *soneme* which specifically refers to the timbral features of the elementary units of music. This term is inspired from a terminology proposed by Vecchione further described in section 10.2.

10.1.3 Phonemes and notes as they are produced

Speech and music may also be compared from a purely acoustical point of view, as does Wolf in a recent article [157].

Acoustical feature	Music	Speech
Fundamental frequency (when quasi periodic)	*pitch component of melody* categorised notated precision possible	*pitch component of prosody* not categorised not notated variability common
Temporal regularities and quantisation on a longer time scale	*rhythmic component of melody* categorised notated precision possible	*rhythmic component of prosody* not categorised not notated variability common
Short silences	*articulation* sometimes notated	*parts of plosive phonemes* implicitly notated
Steady formants	*components of instrumental timbre* not notated not categorised	*components of sustained phonemes* notated categorised
Varying formants	*not widely used* —	*components of plosive phonemes* categorised notated
Transient spectral details	*components of timbre* not categorised sometimes notated	*components of consonants* categorised notated

Fig. 10.2 Some acoustical features of music and speech signals [157].

In the table he compiled (Fig. 10.2), music and speech are compared on the basis of acoustical features such as fundamental frequency, temporal regularities, short silences, steady formants, varying formants, transient spectral details. Wolf indicates, for example, that steady formants are components of sustained phonemes, which are notated and categorized, and of instrumental timbre, which are neither notated nor categorized.

As described in Chapter 3, notation systems have been developed for plucking positions and plucking techniques. Western systems, such as Company's system, propose indirect notations of timbres since only the techniques to achieve these timbres are notated. In the notation system for the Chinese lute, symbols which are pronounced as one-syllable sounds (such as "Kou") remind the performer as to how a particular timbre is produced, just as

the symbol “a” reminds the reader about how to pronounce the speech sound (open and round mouth, low tongue, etc.)

The table also indicates that “Varying formants”, which are components of plosive phonemes in speech, are *not widely used* in music. In light of the listening test reported in Chapter 9 where participants spontaneously associated [b], [k] and other plosives to the attack portion of guitar tones, we can say that the attack of plucked-string tones plays the role of plosive phonemes, although no varying formants are involved. This suggests that it is limiting to compare speech and music solely on the basis of their acoustical features. Rather than a production-oriented perspective, we propose to adopt a perceptual point of view. Though acoustical analogies between instrumental sounds and speech sounds are not systematic, they sufficiently enable instrumental music to give the illusion of speech.

10.2 *Sonetics* and *sonemics*

10.2.1 Definition

The musicologist Bernard Vecchione proposes to draw systematic parallels between the disciplines studying speech and music. While phonetics and phonemics examine the nature and the function of speech sounds, *sonetics* and *sonemics* would be the disciplines devoted to the study of the musical sounds.

According to Vecchione, *sonetics* is a subdomain of computer research applied to music and acoustics that is devoted to sound analysis and synthesis models, to the study of perceptive functions involved in music listening, and to the characteristics of sound-producing gestures [163]. Vecchione specifies that the studies forming the backbone of this discipline aim to:

- analyse the acoustical signals of musical performance,
- establish precise relations between acoustical signals and characteristics of the signal-producing gesture,
- identify regular associations between certain types of acoustical signals and perceptual dimensions of timbre.

By analogy with the distinction between *phonetics* (the scientific study of the sounds of language and of the spoken communication process) and *phonemics* (the study of the

function of phonemes in a given language), *sonemics* can be defined as the science of the functional classification of acoustical units related to either gesture (production) or reception. Because acoustically different units can be perceived as equivalent in their meaning or function, the two disciplines are complementary. *Sonetics* studies music in its acoustical, psychoacoustical and gestural reality and *sonemics* studies the cognitive activities involved in the production and reception of music [162].

10.2.2 Overlooked: the prosody of language and the sonemes of music

Speaking is a common activity in which all people participate. Since speech conveys information, the precision of its constituents is crucial. For speech sounds, correlations between perceptive features and articulatory features have been established for a long time. However, the study of prosody and paralanguage (the music of language) in general has been attributed much less attention.

Since music performance is a very specialized activity (only a fraction of a population learn how to play an instrument), its scientific study has generated a much smaller body of research, in comparison with the field of linguistics.

The study of the control of timbre by professional musicians is even more specialized. At a beginner level, the musician is only concerned with producing tones with in correct pitch and rhythm. It is only with further musical training that the musician becomes concerned with refining articulation to achieve subtle variations in timbre.

While the acoustics of musical instruments and psychoacoustics are well-established fields (whose research is reported not only in scientific articles and papers but also in books), the scientific study of how a performer manipulates an instrument to obtain particular timbres – what Vecchione calls sonetics – has not yet been established as a separate field, though it bridges the knowledge between acoustics and psychoacoutics.

10.3 Drawing parallels between speech and instrumental music

Here are the parallels we propose to draw between the elementary units of speech and music, as well as the disciplines dedicated to their study.

The first section of Table 10.1 presents disciplines studying aspects of speech and music.

- **Anatomy** (from Greek *anatomê*, "dissection") is the scientific study of the shape,

SPEECH	INSTRUMENTAL MUSIC
anatomy physiology phonetics (articulatory, acoustic, auditory) phonemics (or phonology)	organology mechanics/acoustics sonetics (gestural, acoustic, auditory) sonemics (or sonology)
phoneme phone allophone diphone	soneme sone allosone disone
consonant vowel	attack, transient (harmonic) sustain or release
prosody	pitch contour and rhythm
text	score
phonemic system	sonemic system

Table 10.1 Parallels between disciplines studying aspects of speech and music, between elementary units, modulation and notation of speech and music.

the disposition and the structure of organs.

- **Organology** (from Greek *organon*, meaning a "tool" or "instrument" used in some activity or trade) is the study of musical instruments. It embraces study of instruments' history, instruments used in different cultures, technical aspects of how instruments produce sound, and musical instrument classification.

- **Physiology** is the scientific study of the normal functionning of a living organism or of its parts.

- **Acoustics** (from Greek *akouein*, "to hear") is a subfield of mechanics studying sounds.

- **Phonetics** is the scientific study of the sounds of language and of the spoken communication process. Phoneticists are more concerned with the sounds of speech than the symbols used to represent them. Phonetics has three main branches:
  - **articulatory phonetics** is concerned with the positions and movements of the lips, tongue, and other speech organs in producing speech;

  - **acoustic phonetics** is concerned with the acoustic properties of the speech sound;
  - **auditory phonetics** is concerned with speech perception.

- **Sonetics** is the study of music in its gestural, acoustical and psychoacoustical reality:
  - **articulatory sonetics** is concerned with the relations between acoustical signals and characteristics of the signal-producing gesture;
  - **acoustic sonetics** is concerned with the acoustic properties of the musical sound;
  - **auditory sonetics** is concerned with timbre perception.

- **Phonemics** (or phonology) is the study of the function of phonemes in a given language and the opposition and contrasting relations in the system formed by the sounds of this language.

- **Sonemics** (or sonology) is the science of functional classification of acoustical units related to either gesture (production) or reception.

The second and third sections of Table 10.1 present the elementary units of speech and music.

- **Phoneme** (from Greek *phônê*, "voice"): the continuum of all observed speech sounds in a language reduces to a relatively small number of functional contrasts or phonemic classes, called phonemes. Every phoneme contrasts with every other phoneme. Every speech sound is a member of one (and only one) phonemic *class*. As a class, it does not exist but it is the set of all the sounds that it represents. Together all the phonemes include every sound heard in the language. Each phoneme within a phonemic system has its own symbol. The same symbol will not necessarily mean the same thing from language to language, but its significance in each language will be carefully explained. Some languages reveal relatively few phonemes, while others use up to sixty. Most of the languages that have been analyzed employ about thirty-five phonemes.

- **Phones** are the objects of study in phonetics; the phones are the actual speech sounds as uttered by human beings; only phones have an objective existence.

- **Allophones** are phonetically similar phones, that can be grouped in phonemic classes. The phonemes are the centres of classes of allophones; finding these centres is the primary role of phonemics. Allophones of the same phoneme never contrast; they are always in complementary distribution or in free variation with one another. For example, *p* as in *pin* and *p* as in *spin* are allophones in the Engligh language. The symbol for a phoneme is the symbol for all its allophones.

- **Diphones** are pairs of consecutive phones, such as [ba].

- **Soneme** (from Latin *sonus*, "sound"): sound element in a musical instrument "language". The soneme could be defined as an element in the palette of timbre nuances achievable on a given instrument, labelled with verbal descriptors such as round, dark, nasal, hollow, etc.

- **Sones** are the objects of study in sonetics; the sones are the actual musical sounds as produced by performers; only the sones have an objective existence.

- **Allosones**: on a given instrument, there can be more than one way to produce a sound that can be qualified as round, for example. The allosones would be all the possible instances of allosonic classes whose centres are sonemes. Furthermore, in the vocabulary used by musicians to describe timbre, there are many synonyms. This implies that the vocabulary may be reduced to a smaller number of functionally contrasting qualifiers. For example, radiant is a synonym of luminous, rich is a synonym of full-bodied and precise is a synonym of focused.

- **Disones** are pairs of consecutive sones. For the case of most traditional instruments, the equivalent of a **consonant** is the **transient** (most often the attack) and the equivalent of a **vowel** is the **harmonic** portion (**sustain** or **release**) of the instrumental sound. Guitar sounds are usually perceived as **disones** composed of a consonant-like sone followed by a vowel-like sone.

Finally, from an acoustical point of view, we could say that **prosody** is to language what melody (**pitch** and **rhythm**) is to music. In phonetics, prosody[1] is the study of intonation,

[1] In music, the term has a different meaning: it is the study of the concordance rules between the accents of a text and the strong or weak accents of the music that accompanies the text.

accentuation, pitch and rhythm, pauses and duration of phonemes. While prosody is not notated, melody is notated in the form of a **score**. While written languages constitute sophisticated notation systems for spoken languages, there is no such standard notation system for instrumental timbre.

- **Text**: a written language can be considered as a timbral notation that represents the arbitrary collection of sounds the language has chosen to convey its meanings. In the case of Western languages, pitch is left to the discretion of the speaker or reader.
- **Score**: because musical instruments (especially in Western cultures) have been designed to excel in the pitch domain of music, a notation of instrumental music has developed to a high degree of sophistication in this area, together with duration and dynamics, and the timbral quality regarded with secondary importance.

10.4 Applications of a sonemic system

10.4.1 Expression and meaning

In order to be useful and meaningful, language as a "culturally tempered system of arbitrary, recurrent, and structured sounds" seems to require a minimum amount of variety. The same may be said about the expressive language of a musical performance. In fact, the difference between a poor and a great guitar performer is that the poor performer is not "articulated enough" and is not able to "make the guitar sing".

10.4.2 Perceptive descriptions of sounds

Before significant advances were made in physical anatomy and in knowledge of the workings of the body, descriptions of speech sounds were also perceptive: vowels were <u>bright</u>, <u>closed</u>, <u>stuffy</u>, etc. [158]. Then, articulatory phonetics developed as a complete discipline aiming to study speech sounds at their source. Rather than describing the vowel [u] as <u>closed</u> and <u>dark</u>, phonetical analysis identifies the articulatory parameters of the vowel: [u] is tongue high, tongue back, and lips rounded.

Currently, in Western cultures, the description of instrumental timbre has not evolved beyond the use perceptive terms: timbres are <u>round</u>, <u>dark</u>, <u>bright</u>, <u>velvety</u>, etc.

10.4.3 Learning a language

Once the phonemic system of a language has been deduced, it can be used in a number of ways. Knowledge of the phonemic system of a language can greatly facilitate the learning of that language (especially if it is not the mother tongue).

Applied to the practice of a musical instrument, knowledge of the sonemic system of an instrumental language should facilitate the learning of that language.

Guitarists use verbal descriptors to describe timbres (equivalent of diphones) but the vocabulary is sometimes too abstract, leading to misunderstandings between teachers and students, for example. Schneider proposes a guide for altering the timbre of a guitar tone rationally, as opposed to intuitively [30]. He calls it "The rational method of tone production". Schneider comments: "If the guitarist is aware of each of the timbral parameters that define the tone and is able to relate these parameters to the mechanical processes of the instrument and to his own actions, the player can change colors at will rather than by chance". Vennard also recommends an objective pedagogy: "A knowledge of the mechanism is the foundation of an objective pedagogy, and a mastery of the technic is the prerequisite for artistic expression" [147] (p. 220).

10.4.4 Orthography and notation

The phonemic system of a language provides a basis for the development of an orthography for the language. Applied to music, the sonemic system of an instrumental language can provide a basis for the development of a notation for expressive timbre nuances.

The American composer Henry Cowell noted with regrets the absence of such a system and its implication on the authentic reproduction of various repertoires: "Since there is no notation of tone-quality, a tradition has grown as to how the tone should be played in Chopin, Debussy, and others; but tradition is a vague thing and is subject to subtle alterations. Chopin and Debussy might be better performed if they had been able to write down the exact shades of tonal values they desired in their works" [159] (pp. 34-35).

A few attempts to develop notation systems have been made. The composer Donald Martino [73] developed a symbolic notation system that uses phonetic models for differentiating the attacks of wind instruments, the instrumental technique borrowing from the vocal [30].

10.4.5 Comparing languages

The phonemic system of a language forms a basis for further analysis of the language on more complex levels; it forms a basis for comparing this language with other languages.

The sonemic system of an instrumental language can form the basis for comparing this language with other instrumental languages. It would be particularly useful in orchestration, when having to combine the timbres of different instruments of the orchestra. It could also be useful when comparing the timbres of different guitars varying in their structure and material.

10.5 Parallels between guitar tones and speech sounds

10.5.1 Interdependance between phones and sones

In speech, phones are combined into diphones. Many combinations of phones are possible, such as a consonant followed by a vowel: [ba], [da], [ga], [be], [de], [ge], ... With the guitar, a particular type of attack has an effect on the release part of the tone since it constitutes the excitation of the tone. Because of this constraint between the two sones of a guitar tone, the guitar is perceived as a voice that can only produce certain syllables. For example, if the attack is sharp as a [k], then the release is brighter, evoking a more acute vowel.

10.5.2 Articulation in speech and guitar

The same sound [t] may be produced by various arrangements of the articulators: the tip of the tongue may be placed anywhere from the point of the upper teeth to the soft palate, and the resulting voiceless stop will more or less resemble most people's concept of what an ideal [t] should sound like.

Similarly, a whole set of gestures will achieve similar timbres and the playing technique employed to achieve one particular timbre might also vary among players according to their experience, the size and shape of their arms, hands and fingers, and the softness/hardness of their nails For example, one cannot say that playing to the right (as recommended by Tarrega) is more favourable then playing to the left (with the hand in the same axis as the forearm).

Part III

Guitar Timbre and Gesture Parameter Extraction

Chapter 11

Indirect Acquisition of Plucking Position

Contents

This chapter addresses the problem of extracting the plucking point information from a recording. This work is the continuation of the research I accomplished at the Center for Computer Research in Music and Acoustics, Stanford University, under the supervision of Prof. Julius O. Smith in partial fulfillment of an Engineer degree [48]. The results of this previous research project are summarized in section 11.3 in order to situate the new research within context. The new method proposed in this thesis for the estimation of the plucking position uses an iterative weighted least-square algorithm.

11.1 Indirect acquisition of instrumental gesture parameters

The *indirect acquisition* of an instrumental gesture parameter consists in capturing the characteristics of instrumental gesture by analyzing of the acoustical signal, namely from a recording [70]. This differs from the *direct acquisition* performed with sensors on the instrument or on the performer. In recent years, there has been an important development of technologies related to sensors and gestural interfaces. For example, many musical instruments can be augmented with devices that can monitor the performer's actions (choice of keys, pressure applied to a mouthpiece, etc.) and turn it into MIDI information.

Direct acquisition is clearly a simpler way to capture the physical features of a gesture, but it is potentially invasive and may ignore the interdependency of the different variables. For example, sensors on a clarinet detect the air jet speed and the fingering but do not account for the coupling between the excitation and the resonator. As opposed to direct acquisition, indirect acquisition is based on the assumption that the performance parameters can be extracted from the signal analysis of the sound produced by an instrument. The main difficulty of this task is to determine in the signal, the specific acoustic signature of a particular performance parameter that has a perceivable influence on the sound.

The data consists in the recording of musicians playing tones with specific gestures, attempting to vary one gesture parameter at a time.

In the first stage of the analysis of the data, basic sound parameters are extracted from the acoustic signal through time- and frequency-domain analysis. These low-level parameters include [74]:

- the short-time energy (related to the dynamic profile of the signal),

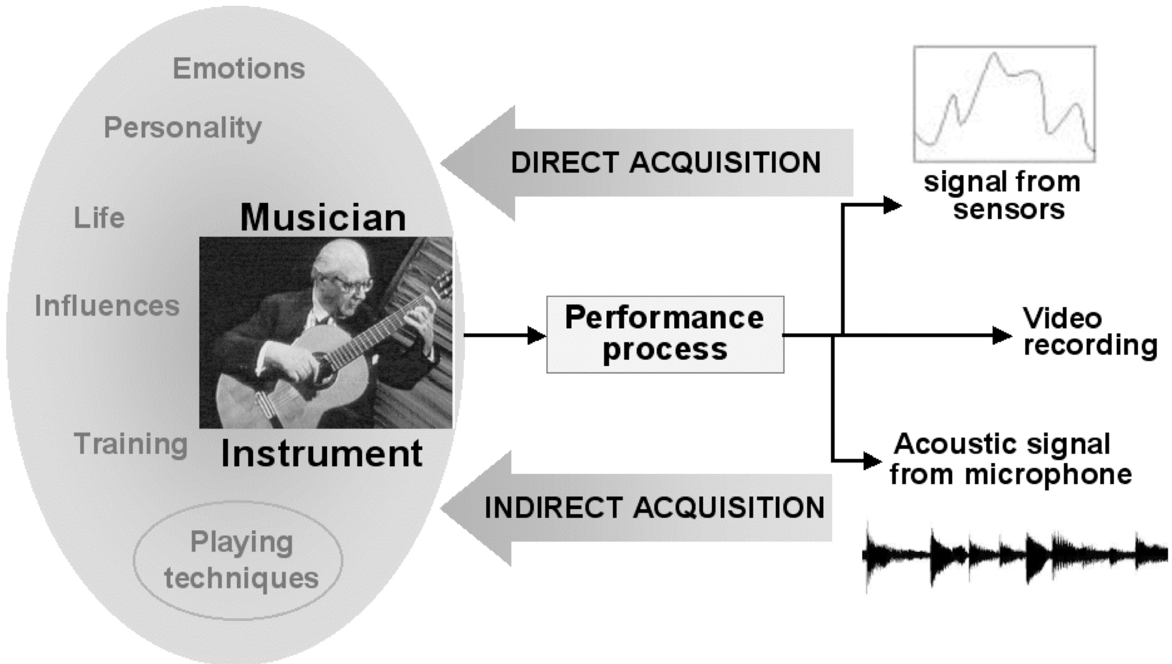


Fig. 11.1 Direct vs indirect acquisition of instrumental gesture parameters.

- the fundamental frequency (related to the sound melodic profile),
- the spectral envelope (in particular, the location of the resonances in the spectrum),
- the amplitudes, frequencies and phases of sound partials, and
- the power spectral density.

With the knowledge of physical mechanisms occuring in musical instruments, physical model parameters can be derived from the basic sound parameters. These parameters generally allow direct access to the instrumental gesture parameters.

In our study of the guitar timbre, the impact of the variation of instrumental gesture parameters on the perceived timbre was described in Chapters 7 and 8.

Although this study addresses issues related to the general problem of timbre recognition, the approach that we propose for the analysis of instrumental timbre differs from the phenomenological approach taken in many timbre recognition systems described throughout literature [77, 82, 84]. Timbre recognition systems implementing neural networks or using Principal Component Analysis require a learning stage, meaning that a timbre can

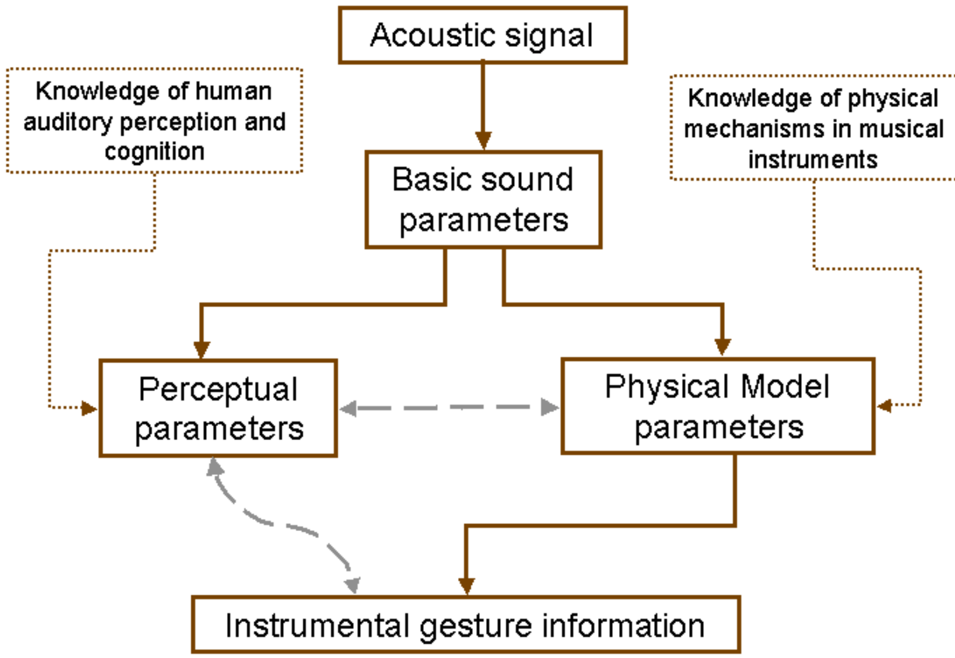


Fig. 11.2 From acoustic signal to gestural information.

only be identified and labelled after being compared to other typical examples of that timbre. Therefore, they do not make explicit the relationships between the physical phenomena, the performer's actions and the obtained timbre. Here, we rather propose to develop analysis tools that use the knowledge of the physical phenomenon occuring in the musical instrument and its effect on the acoustical signal, leading to an analytical model of the interaction between the performer and the instrument.

11.2 Indirection acquisition of plucking position

11.2.1 Effect of plucking position on magnitude spectrum

Varying the plucking location greatly affects the spectrum of the sound, similar to the effect of a comb filter, which manifests itself by the presence of equally spaced attenuations (zeroes) in the spectral envelope [6].

As shown in Chapter 4, the amplitude $C_y[n]$ of the nth mode of the displacement of an ideal vibrating string of length l plucked at a distance p from the bridge with an initial

vertical displacement h is given by :

$$C_y[n] = \frac{2h}{n^2\pi^2 R(1-R)} |\sin(n\pi R)| \tag{11.1}$$

where $R = p/l$ is the relative plucking position, defined as the fraction of the string length from the point where the string was plucked to the bridge.

In this context, since we are interested in the extraction of the parameter R, we will use the notation $C_n(h, R)$ (rather than $C_y[n]$) which expresses the coefficient as a function of two parameters, the relative plucking position R and the height of the initial displacement h.

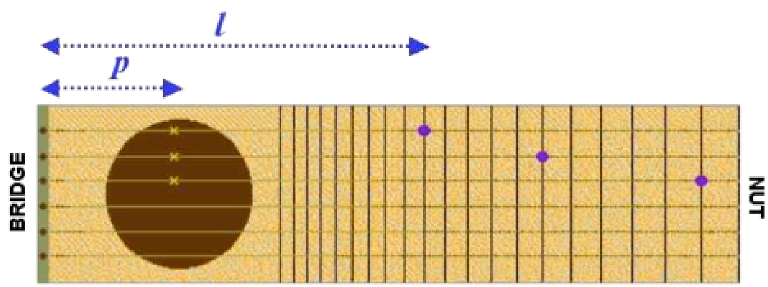


Fig. 11.3 Plucking point at distance p from the bridge and fingering point at distance l from the bridge on a guitar neck.

11.2.2 Pratical limitations to the estimation of the plucking position

In particular circumstances, the output from the string (force at the bridge) lacks the harmonics that have a node at the plucking point. A simple way of estimating the plucking point location along the string from a recording is to pinpoint the missing harmonics in the spectrum ($C_n = 0$). However, the string is not usually plucked exactly at a node of any of the lowest harmonics. Since the amplitude of the higher harmonics is considerably smaller, it is not always possible to accurately detect the plucking point by simply searching for the missing harmonics in the magnitude spectrum.

Fig. 11.4 illustrates how the spectral envelope is sampled to obtain the spectrum corresponding to a given relative plucking position R. On the left, $R = 1/5$, and the spectrum has zeroes (harmonics of order 5 and its integer multiples are cancelled). On the right, $R = 0.234$, and although the spectral envelope has zeroes, the sampling of the spectral envelope is such that the harmonics do not fall on those null frequencies.

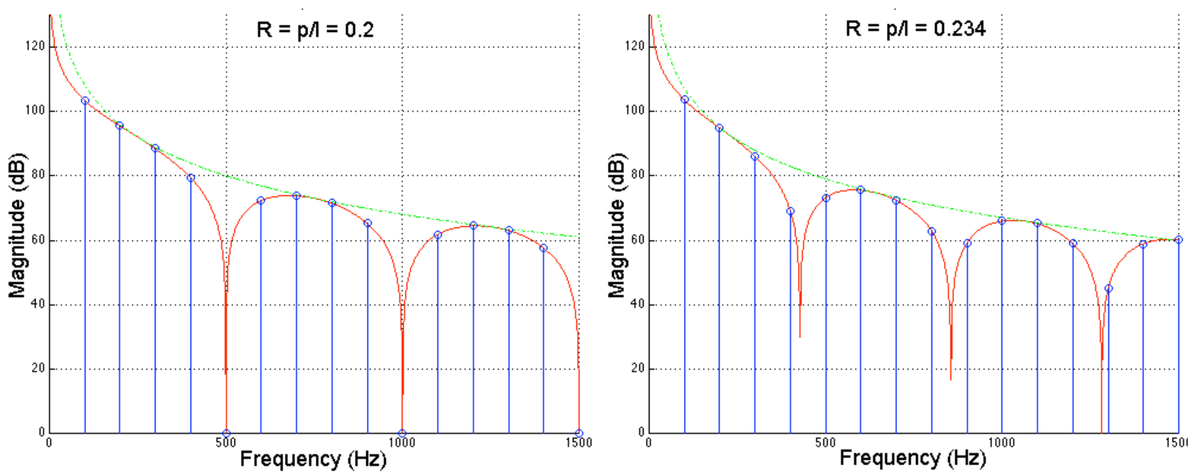


Fig. 11.4 Ideal string spectra for $R = 1/5 = 0.2$ on the left and for $R = 0.234$ on the right. In the first case (R is the inverse of an integer), some harmonics are missing. In the second, the sampling of the spectral envelope is such that none of the harmonics are missing.

Estimation of the plucking position from a recorded sound is an intrinsically difficult problem since a recorded tone can include contributions of several delays of approximately the same magnitude, such as early reflections from objects near the player, the floor, the ceiling, or a wall. The guitar body can also induce significant filtering. Therefore, recordings conditions should be carefully set.

Another practical problem may arise from the nonlinear properties of the string. More specifically, the amplitude of vibration of a weak harmonic can gain energy from other modes so that its amplitude begins to rise, reaching a maximum about 1000 ms after the attack, and then begins to decay [11]. This is often seen in the analysis through time of the harmonic envelopes of guitar tones. The non-linear properties of the string may result in a preference for a time-domain approach using, for example, the short-term autocorrelation function.

11.2.3 Review of plucking point estimation methods

In [83], three analysis techniques were used to investigate four instrumental gesture parameters of the guitar (finger position along the string, inclination between finger and string, inclination between hand and string, and degree of relaxation of plucking finger). Among

these analysis techniques, Principal Component Analysis is used to verify that each of the instrumental gesture parameters induces significant changes in the cepstral envelope. However, it is not clear whether this methodology constitutes an indirect acquisition system since the four sets of guitar tones were analyzed separately.

A time-domain approach for estimating the plucking point is proposed by Välimäki & Penttinen in [43]. It is not an indirect acquisition system per se since it uses an under-saddle pickup. The algorithm is based on investigating the time lag between two consecutive pulses arriving at the bridge of the guitar. The method determines the minimum of the autocorrelation function for one period of the signal.

A frequency-domain approach is proposed by Bradley & al. in [33]. The plucking position is determined from the data by finding the value of the relative plucking position R that minimizes the absolute value of the error between the ideal string spectrum and the sampled-data spectrum. An improved implementation of the method suggested by Bradley & al. is reported in [48] (Engineer thesis of the author at CCRMA) and [49]. A summary of the results of this research is presented in the next section.

11.3 A frequency-domain method for extracting plucking position

11.3.1 Description of the method

Fig. 11.5 summarizes our implementation (reported in [48]) of the method proposed by Bradley & al. in [33]. Our implementation includes supplementary units that render possible the automatic processing of the audio recording of a performance. Here is the description of the function of the different units.

Attack Detection: The energy for successive blocks of 512 samples is calculated while an increase of the energy by a factor of 2 turns on a flag. If the energy increases by a factor of 2 two or more times in a row, successive alarms have to be eliminated. After the beginning of each tone is identified, a section of the sampled waveform is chosen for analysis. The starting sample of the section is chosen at approximately 1/8th of the distance in samples between two attacks. This roughly corresponds to the beginning of the stationary part of the sound.

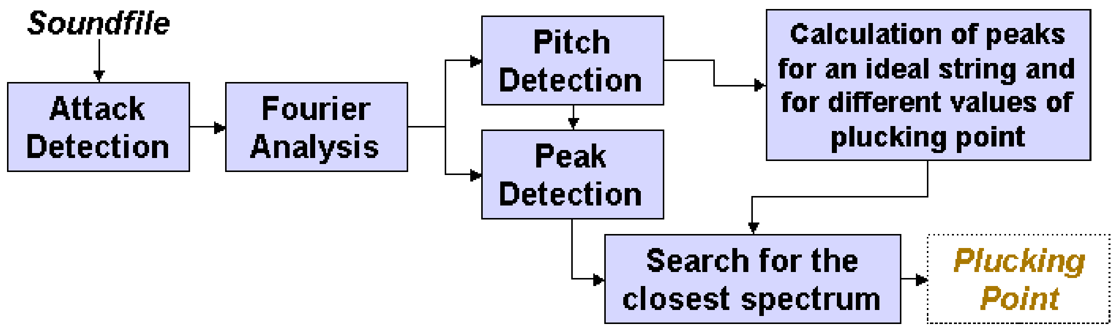


Fig. 11.5 Block-diagram for estimation of the plucking point [48].

Fourier Analysis : The spectrum is generated by windowing the waveform and performing a longer Fast Fourier Transform (the number of bins chosen so that two overtone peaks in the spectrum will not overlap). 2^{12} (4096) samples from the sound file are extracted from the middle of the tone (after the attack), starting at the index provided by the **Attack detection** unit. The sound portion is windowed with a Hamming window then the FFT is computed with a zeropadding factor of 6 and a parabolic interpolation.

Pitch Detection: The fundamental frequency is determined by finding the first maximum in the autocorrelation function (occuring at the fundamental period) [64].

Peak Detection: In this unit, the harmonics are identified. With the pitch value determined by the **Pitch Detection** unit, we look for a maximum in narrow intervals around integer multiples of the fundamental frequency (Fig. 11.6).

Plucking point estimation : The plucking position is determined from the data by finding the value of the relative plucking position R that minimizes the absolute value of the error between the ideal string spectrum and the sampled-data spectrum, as expressed by Eq. (11.2), where H_n is the measured set of sampled string harmonic information.

$$\varepsilon = \sum_{n=1}^{N} \left| |H_n| - \left| \frac{2h}{n^2 \pi^2 R(1-R)} \sin(n\pi R) \right| \right| \tag{11.2}$$

An error surface for various values of R is constructed by evaluating the error criterion ε for various values of R; the minimum of the error indicates an estimation of the plucking

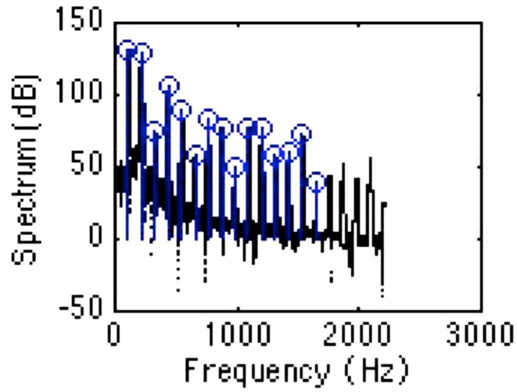


Fig. 11.6 Spectrum and peak detection.

position, as illustrated on Fig. 11.7.

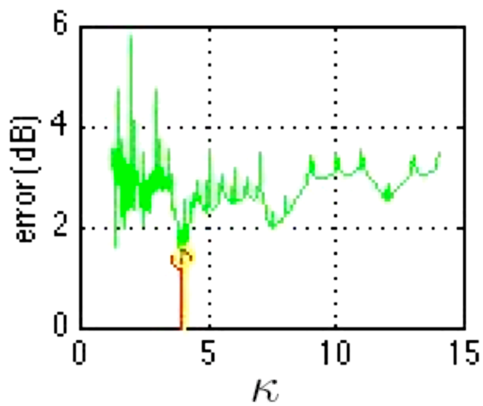


Fig. 11.7 Error surface for various values of relative plucking position. The minimum of the error is chosen as the plucking position. The horizontal axis on this graph is the inverse of the relative plucking position ($\kappa = 1/R$). For example, if the string is plucked at a third of its length, $\kappa = 3$ [48].

11.3.2 Results

Fig. 11.8 displays the results of the analysis for four plucking positions from the bridge (12, 13, 14 and 15 cm). The estimations are 12.2, 13.1, 14.5 and 14.6 cm respectively. On the figures, the left window shows the Fourier analysis of a 4096-sample portion of the sound with peak detection indicated by circles. The central window shows the error curves for

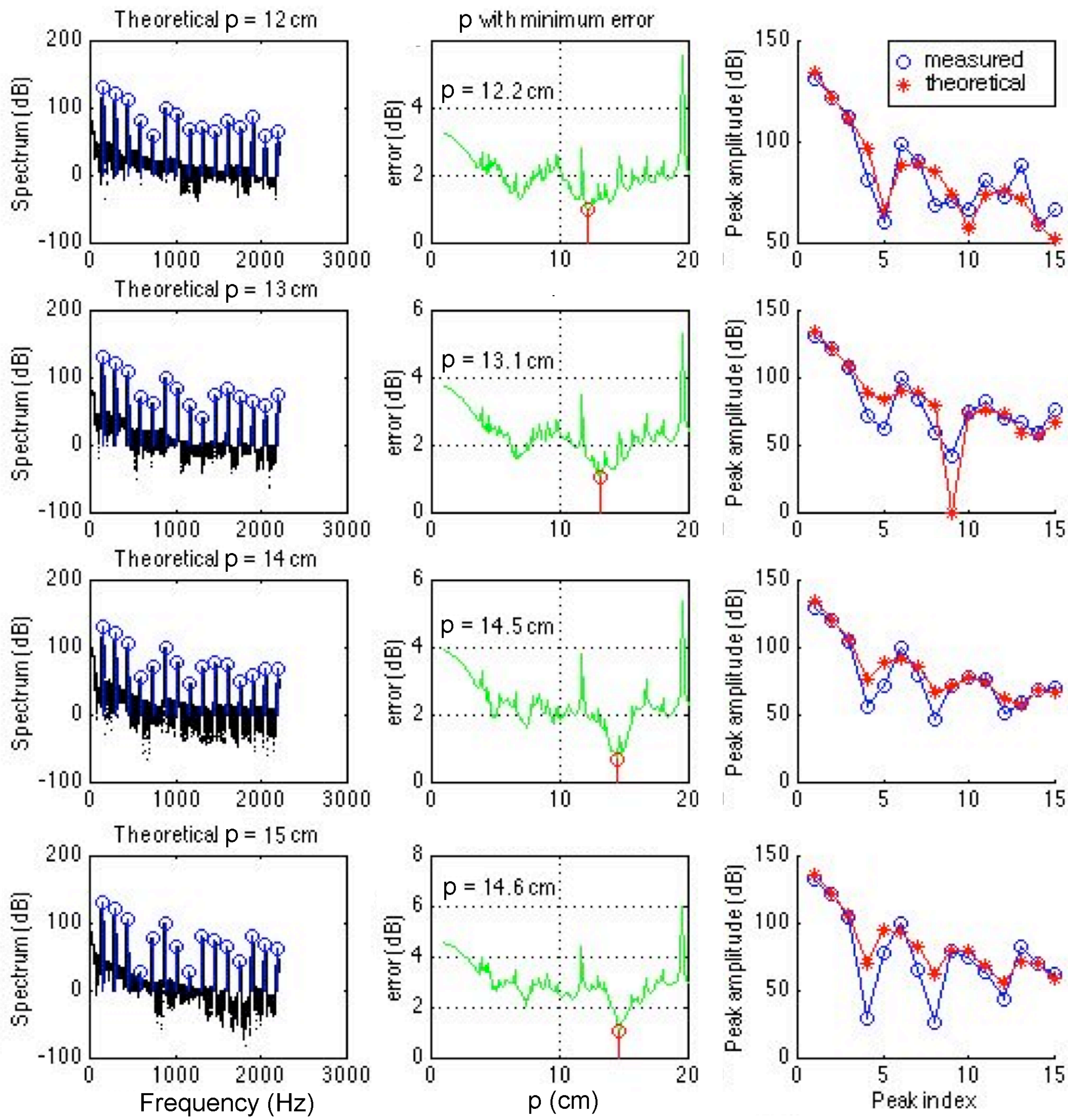


Fig. 11.8 Plucking position estimation for tones played on the open D-string of a classical guitar with plucking position from the bridge = 12, 13, 14 and 15 cm [48].

various values of the absolute plucking position p ranging from 1 to 20 cm. The minimum is indicated by a circle and the corresponding p value is displayed. The right window is the comparative display of the detected peaks (o) and of the ideal string spectrum ($*$) based on the intended plucking position.

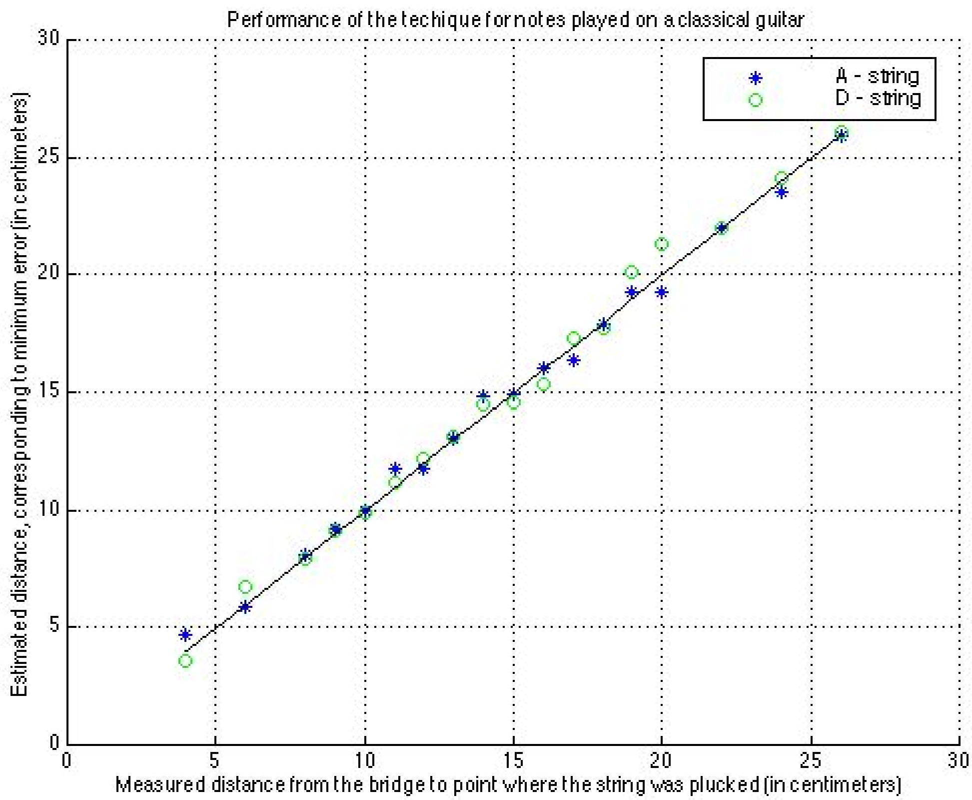


Fig. 11.9 Plot summarizing the results for 18 plucks on open D- and A-strings of the classical guitar [48].

Fig. 11.9 summarizes the results obtained for the 18 plucking points on the open A-string and open D-string. The graph displays the estimated distance versus the measured distance on the string when the tone was played. The margin of error was less than 1 cm.

Although this method presents a satisfying accuracy, it is computationally heavy since a large number of theoretical spectra have to be calculated and compared to the observed spectrum. Moreover, this method implies a quantification of the plucking position value, leading to rounding errors. The new method presented in the next section is more direct

and computationally efficient.

11.3.3 Information on sound data base

The recorded tones examined were played with a triangular shaped plastic pick, 0.88 mm in thickness, on a plywood classical guitar strung with nylon and nylon-wrapped steel Alvarez strings. The intended plucking locations were precisely measured and indicated on the string with a marker. The tones were recorded with a Shure KSM32 microphone in a sound-deadened room, onto digital audio tape at 44.1 kHz, 16 bits. The microphone was placed in front of the sound hole, approximately 25 cm away; at this distance, a combination of waves emanating from different parts of the string is captured, thereby limiting the filtering effect of the pickup point.

11.4 Extraction of the excitation point location on a string using weighted least-square estimation

This section describes a new method for estimating plucking point location. Starting from a measure related to the autocorrelation of the signal as a first approximation, a weighted least-square estimation is used to refine the comb filter delay value to better fit the measured spectral envelope. The general procedure is illustrated in Fig. 11.10.

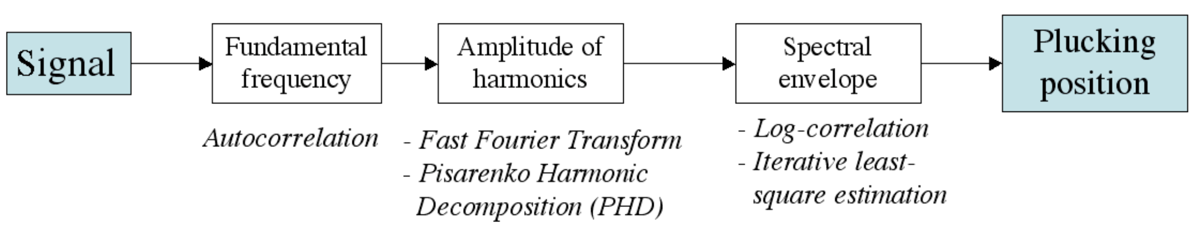


Fig. 11.10 Block-diagram of general procedure from the acoustic signal to the plucking position.

For determining the magnitude of the harmonics, Pisarenko Harmonic Decomposition (PHD) was implemented and compared to Fast Fourier Transform (FFT). This work is reported in [51]. The PHD algorithm was too sensitive to the nature of the background noise (this algorithm works best when the noise is white). Hence, since the PHD algorithm was not more accurate, the FFT was used for the method described in this section.

11.4.1 First approximation for R from Log-Correlation

The autocorrelation function $a(\tau)$ of a periodic signal $x(t)$ with fundamental period T_o can be expressed in terms of its Fourier series magnitude coefficients C_n in the following way (see Appendix A for details):

$$a(\tau) = C_o^2 + \frac{1}{2}\sum_{n=1}^{\infty} C_n^2 \cos\left(\frac{2\pi}{T_o} n\tau\right) \tag{11.3}$$

While the long-term features of the autocorrelation function are very useful for estimating the fundamental frequency of a periodic signal (since it shows a maximum at a lag corresponding to the fundamental period T_o), its short-term evolution reveals information about the plucking position.

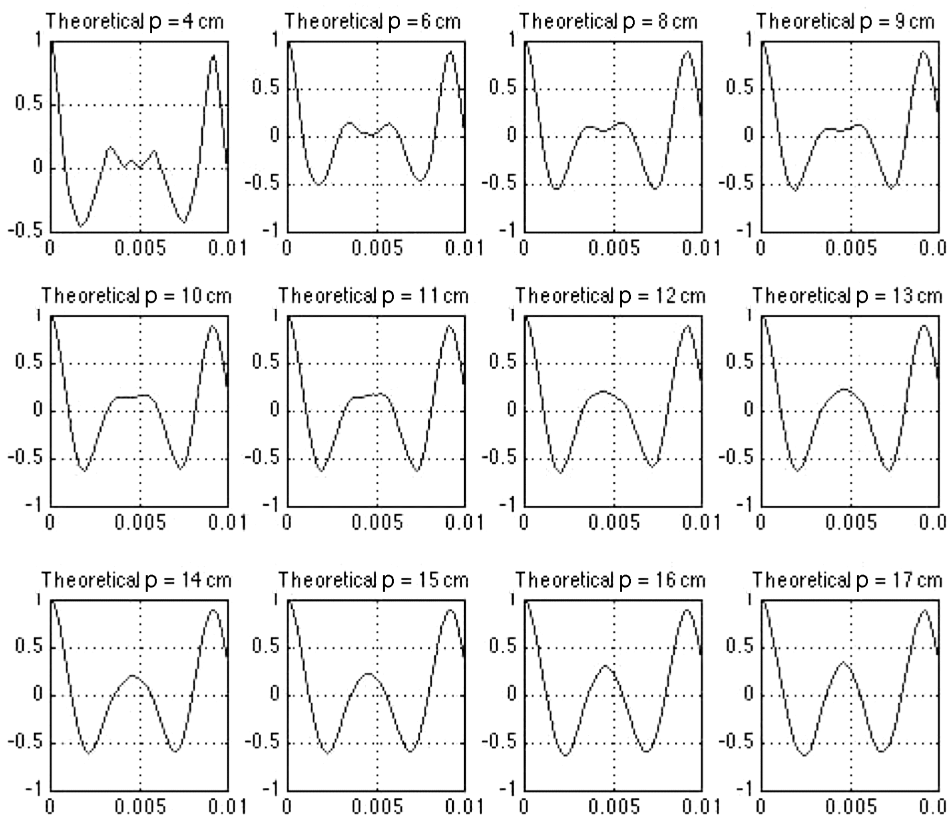


Fig. 11.11 Autocorrelation graphs for 12 guitar tones plucked at distances from the bridge ranging from 4 cm to 17 cm.

Fig. 11.11 displays the plots of the autocorrelation function calculated for 12 recorded

guitar tones plucked at various distances from the bridge on an open A-string (fundamental frequency = 110 Hz). As expected, the graphs show a maximum around $1/110 = 0.009$ seconds, the fundamental lag of the autocorrelation. One can also see that the autocorrelation takes different shapes for different plucking positions, but the information about the comb filter delay can not be extracted directly from these graphs. In order to detect the low amplitude harmonics, we modify the structure of the autocorrelation function by taking the log of the square of the Fourier coefficients (and by dropping the DC component). This emphasizes the contribution of low amplitude harmonics (around the *valleys* in the comb filter frequency response) by introducing large negative weighting coefficients. The obtained *log-correlation* is expressed as follows:

$$\Gamma(\tau) = \sum_{n=1}^{N} \log(C_n^2) \cos\left(\frac{2\pi}{T_o} n\tau\right) \tag{11.4}$$

Fig. 11.12 displays the log-correlation graphs for the same 12 recorded guitar tones (as for Fig. 11.11). As expected, the log-correlation plots reveal an interesting pattern: the global minimum appears around the location of the lag corresponding to the plucking position. Therefore, it can be concluded that the relative plucking position can be approximated by the ratio

$$R \approx \frac{\tau_{min}}{\tau_o} \tag{11.5}$$

where τ_{min} is the lag corresponding to the global minimum in the first half of the log-correlation period, and τ_o is the lag corresponding to the fundamental period T_o, as illustrated on Fig. 11.13.

11.4.2 First approximation for h

A first approximation h_o for the vertical displacement h is also needed in order to initialize the weighted least-square procedure. h_o can be determined from the first approximation R_o of R and the total power of the harmonic components in the observed spectrum $\sum_{n \in I_W} C_n^2$,

$$h_o = R_o(1 - R_o)\frac{\pi}{2}\sqrt{\frac{\sum_{n \in I_W} C_n^2}{\sum_{n \in I_W} \frac{\sin^2(n\pi R_o)}{n^4}}} \tag{11.6}$$

I_W refers to the set of harmonics that are given a significant weight in the second stage

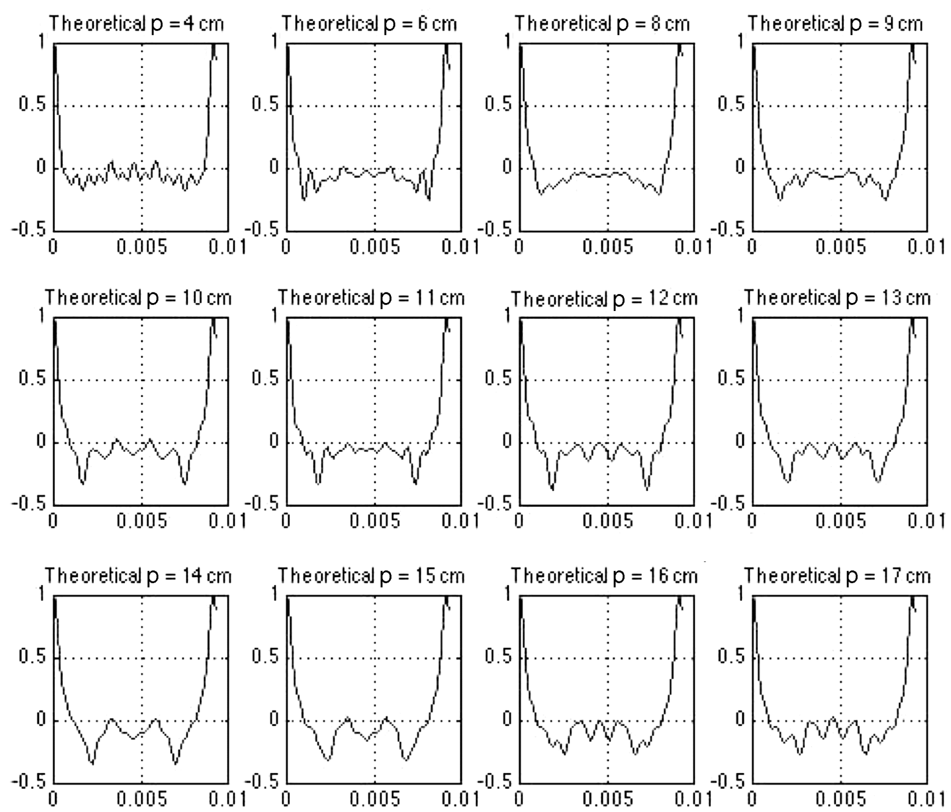


Fig. 11.12 Log-correlation graphs for 12 guitar tones plucked at distances from the bridge ranging from 4 cm to 17 cm.

of the approximation (as described in the next section).

11.4.3 Iterative refinement of R value using weighted least-square estimation

The second stage of the estimation consists in finding the values of h and R that minimize the distance between the theoretical expression of the ideal string magnitude spectrum $\hat{C}_n(h,R)$[1] (Eq. 11.1) and its observation $C_n(h,R)$ in the least-square sense [61].

As illustrated on Fig. 11.14, rather than using the magnitude coefficient $\hat{C}_n$ (whose phase is 0 or π), we use the power coefficients $\hat{C}_n^2$ for which it is not necessary to recover the phase. $\hat{C}_n^2(h,R)$ is proportional to h^2 and $\sin^2(n\pi R)$ and is therefore a non linear expression in terms of h et R. A least-square estimation technique can still be employed after linearizing

[1] $\hat{C}_n$ is considered here to be a model of the amplitude, hence the hat ($\hat{}$) while C_n represents *measured* values or *observed* values.

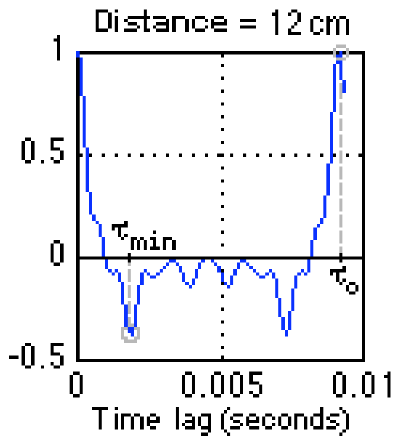


Fig. 11.13 Log-correlation for a guitar tone plucked 12 cm from the bridge on a 58 cm open A-string. Ratio $\frac{\tau_{min}}{\tau_o}$ provides a first approximation for relative plucking position R.

$\hat{C}_n^2(h, R)$ with a first order Taylor's series approximation about a first approximation R_o of R and h_o of the height h of the string displacement. It leads to an expression of

$$\hat{C}_n^2(h, R) = \frac{4h^2}{n^4\pi^4 R^2(1-R)^2}\sin^2(n\pi R) \tag{11.7}$$

as a linear combination of the two correcting values $\Delta h = h - h_o$ and $\Delta R = R - R_o$. The first order Taylor's series for the different factors included in Eq. (11.7) are

$$\frac{1}{R^2} = \frac{1}{R_o^2}\left(1 - \frac{2\Delta R}{R_o}\right) \tag{11.8}$$

$$\frac{1}{(1-R)^2} = \frac{1}{(1-R_o)^2}\left(1 + \frac{2\Delta R}{1-R_o}\right) \tag{11.9}$$

$$\sin^2(n\pi R) = \sin^2(n\pi R_o) + n\pi\sin(2n\pi R_o)\Delta R \tag{11.10}$$

$$h^2 = h_o^2 + 2h_o\Delta h \tag{11.11}$$

By multiplying Eq. (11.8) and (11.9), we obtain the expression for the product

$$\frac{1}{R^2} \times \frac{1}{(1-R)^2}$$

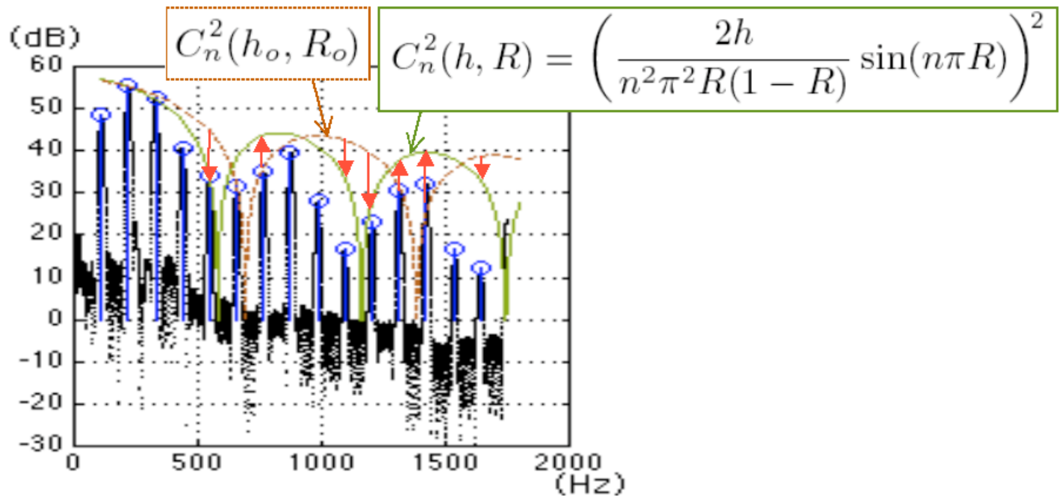


Fig. 11.14 Estimation of spectral envelope in two stages. $\hat{C}_n^2(h_o, R_o)$ is a first approximation. $\hat{C}_n^2(h, R)$ is a better approximation of the spectral envelope based on an iterative weighted least-square estimation.

$$
\begin{aligned}
\frac{1}{R^2(1-R)^2} &= \left[\frac{1}{R_o^2}\left(1 - \frac{2\Delta R}{R_o}\right)\right]\left[\frac{1}{(1-R_o)^2}\left(1 + \frac{2\Delta R}{1-R_o}\right)\right] \\
&= \frac{1}{R_o^2(1-R_o)^2}\left[1 + \frac{2\Delta R}{1-R_o} - \frac{2\Delta R}{R_o} - \frac{4(\Delta R)^2}{R_o(1-R_o)}\right] \\
&= \frac{1}{R_o^2(1-R_o)^2}\left[1 + \frac{2(2R_o-1)}{R_o(1-R_o)}\Delta R\right]
\end{aligned}
$$

after dropping the second order term in $(\Delta R)^2$. By multiplying this last expression by Eq. (11.11) results in

$$
\begin{aligned}
\frac{h^2}{R^2(1-R)^2} &= \frac{h_o^2 + 2h_o\Delta h}{R_o^2(1-R_o)^2}\left[1 + \frac{2(2R_o-1)}{R_o(1-R_o)}\Delta R\right] \\
&= \left[\frac{h_o^2}{R_o^2(1-R_o)^2}\right] + \left[\frac{2h_o}{R_o^2(1-R_o)^2}\right]\Delta h + \left[\frac{2h_o^2(2R_o-1)}{R_o^3(1-R_o)^3}\right]\Delta R
\end{aligned}
$$

after dropping the second order term in $\Delta R \Delta h$. Finally, by multiplying this last expression by Eq. (11.10), we obtain the linearized expression (omitting a $2/\pi^2$ factor):

$$\hat{C}_n^2(h, R) = \hat{C}_n^2(h_o, R_o) + \alpha_n\,\Delta h + \beta_n\,\Delta R \tag{11.12}$$

where

$$\begin{aligned}
\hat{C}_n^2(h_o, R_o) &= \left(\frac{h_o \sin(n\pi R_o)}{n^2 R_o(1-Ro)}\right)^2 \\
\alpha_n &= 2h_o\left(\frac{\sin(n\pi R_o)}{n^2 R_o(1-Ro)}\right)^2 \\
\beta_n &= n\pi\left(\frac{h_o}{n^2 R_o(1-R_o)}\right)^2 \sin(2n\pi R_o) \\
&\quad + \frac{2(2R_o-1)}{R_o(1-R_o)}\left(\frac{h_o \sin(n\pi R_o)}{n^2 R_o(1-Ro)}\right)^2
\end{aligned}$$

Let the difference between the estimated power spectrum nth coefficient and its first approximation be

$$\hat{Y}_n(h, R) = \hat{C}_n^2(h, R) - \hat{C}_n^2(h_o, R_o) \tag{11.13}$$

and the difference between the measured power spectrum nth coefficient and the first approximation be

$$Y_n(h, R) = C_n^2(h, R) - \hat{C}_n^2(h_o, R_o) \tag{11.14}$$

Eq. 11.12 can be expressed as

$$[\hat{C}_n^2(h, R) - \hat{C}_n^2(h_o, R_o)] = \begin{bmatrix} \alpha_n & \beta_n \end{bmatrix} \cdot \begin{bmatrix} \Delta h \\ \Delta R \end{bmatrix} \tag{11.15}$$

which becomes, by grouping the N equations (11.15) for $n = 1, ..., N$, the linear system in matrix form:

$$\hat{Y} = AX \tag{11.16}$$

Since A is a $N \times 2$ matrix, the solution to Eq. (11.16) can be obtained using pseudo-inverse

$$(A^T A)^{-1} A^T$$

or, for better results, its weighted version

$$(A^T W A)^{-1} A^T W$$

where W is a $(N \times N)$ diagonal matrix containing the weights for the least-square errors.

The weighting function can be used to select particular ranges of frequencies or to reject components that are known for deviating from the theoretical comb filter model (near resonant frequencies of the guitar body, for example). A good weighting curve is one that combines a bell curve and a positive sloped ramp. The bell curve increases the contribution of the components in the valleys of the spectrum and the ramp gives more weight to higher order – weaker harmonics – over the whole range of the spectrum.

Finally, the correcting values for h and R are obtained with

$$\begin{bmatrix} \Delta h \\ \Delta R \end{bmatrix} = [\,(A^T W A)^{-1} A^T W\,] \cdot Y$$

minimizing the distance between the model and the observation $||\hat{Y} - Y||$ in a least-square sense. Then, the two parameters R and h are iteratively refined using $h_o + \Delta h$ and $R_o + \Delta R$ as second approximations and so forth.

Between 3 to 10 iterations are generally needed to converge with a criterion error

$$\varepsilon = \left| \frac{R_k - R_{k-1}}{R_k} \right| < 0.001.$$

As expected, the number of iterations decreases with the accuracy of the first approximation. If the first approximation is very rough ($\varepsilon \simeq 0.5$), the number of iterations can increase to about 40, but the algorithm still converges to the right value of R (and h).

Fig. 11.15 displays the plots of the power spectrum of the 12 guitar tones together with the profile of the comb filter before and after iterative refinement. Fig. 11.16 displays the graph of the estimated plucking position $\hat{p}$ vs the actual distance from the bridge p in centimeters for the 12 guitar tones. The diagonal line indicates the target of the estimation (the actual value). The upper window displays a first approximation (obtained with log-correlation for example). The lower window shows the improvement achieved after the refinement of R value using weigthed least-square estimation. For this data set, the average error is 0.78 cm for the first approximation and then is reduced to 0.18 cm after refinement.

11.4.4 Conclusion

We have proposed an efficient method for the extraction of the excitation point location on a guitar string from a recording. It is based on the assumption that the power spectrum of a plucked string tone is comb-filter shaped.

The theoretical expression giving a ideal string magnitude spectrum is proportional to $sin(n\pi R)$ and is therefore non linear. This equation can be linearized with a first order Taylor's series approximation about a first approximation of R and of the height of the string displacement h. These two parameters are then refined iteratively with a least-square estimation technique.

To obtain the first approximation for R, we propose a measure derived from the amplitudes of partials extracted through standard short-time Fourier transform. This measure is a variation on the autocorrelation function for periodic signals which consists in the sum of cosine functions weighed by the log of the square of the Fourier coefficients. We have discussed the properties of this "log-correlation" that emphasizes the minima in the spectral envelope and exhibits a minimum at a lag τ_{min} that provides an estimation of the relative plucking point R by taking the ratio of the *minimum lag* τ_{min} over the *fundamental lag* τ_o.

Many applications can benefit from the algorithm, especially in the context of automatic tablature generation and sound synthesis (extraction of control parameters). This technique can also be used to derive the value of the delay of any kind of comb filter from the spectral peak parameters.

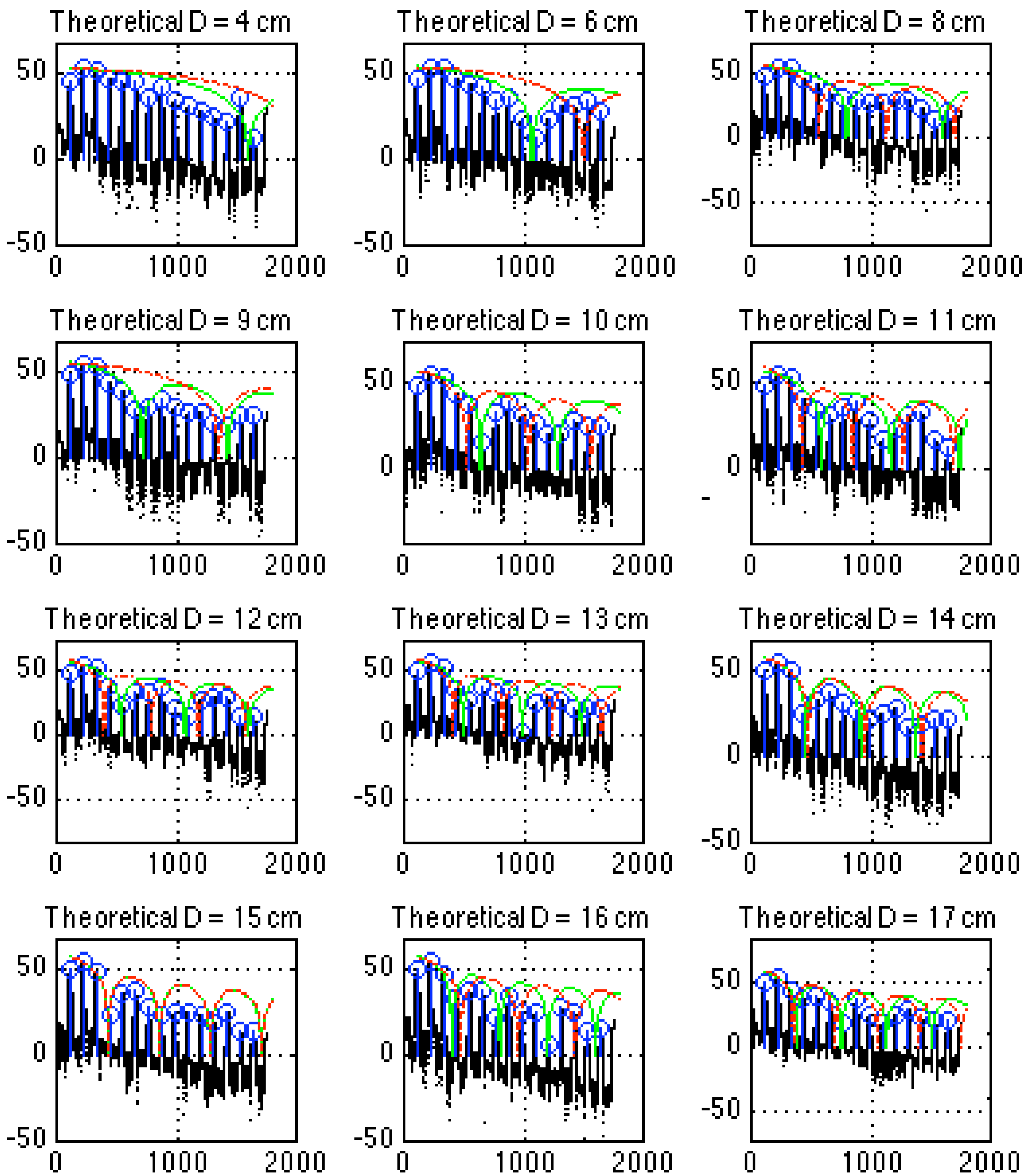


Fig. 11.15 Power spectra of 12 recorded guitar tones with superimposed comb filter model. First approximation plotted with a dark dashed line and final estimation plotted with a light grey line.

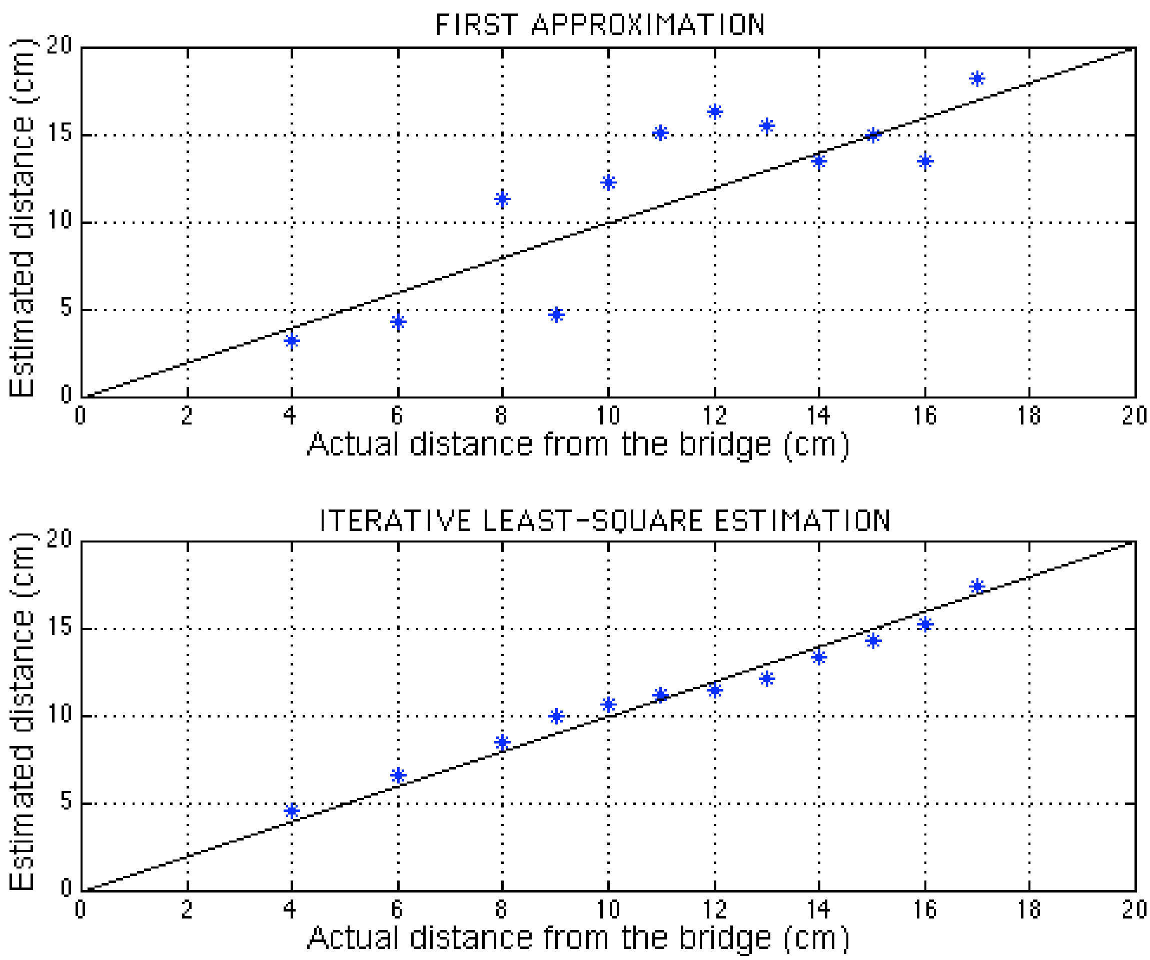


Fig. 11.16 Estimated vs actual plucking distances before (top window) and after (bottom window) refinement of p value using iterative weighted least-square estimation.

Chapter 12

General Conclusion, Applications and Future Directions

Contents

12.1 Conclusion

12.1.1 Different points of view

The classical guitar is an instrument that offers to the skilled performer a vast array of timbral variations. In this thesis, the instrument's timbre was investigated from different perspectives.

From the point of view of the instrument, we identified the static control parameters of timbre, relating to the structural components of the guitar. From the point of view of the performer, we identified the dynamic control parameters of timbre, relating to the gestures applied by the performer on the instrument. For example, by varying the plucking position along the string, the guitarist can control the parameters of the guitar tones' spectral envelope, and modify the perceived timbre. From the point of view of the listener, we explored the rich vocabulary used by guitarists to describe the brightness, the colour, the shape and the texture of the sounds they produce on their instruments. Dark, bright, chocolatey, transparent, muddy, wooly, glassy, buttery, and metallic are just a few of the timbre descriptors that we collected from questionnaires submitted to 22 guitarists. The acoustical basis of this vocabulary was investigated.

12.1.2 Different sources and methodologies

The different points of view called for various sources of information and methodologies.

From physics to signal processing

Since the plucking position is an important parameter of the plucking gesture, a particular attention was attributed to the shape of the spectral envelope induced by this parameter. Starting from the plucked string physical model (obtained from the tranverse wave equation), we derived a digital signal interpretation of the plucking effect which is a comb filter with delay $D = R/f_0$ (relative plucking position over fundamental frequency of the string).

From signal processing to speech perception

Since the vocal quality of the guitar has been remarked upon so often, we searched for formants in the spectral envelope of guitar tones and found what we propose to call "comb

filter formants", centred at frequencies similar to typical vocal formant frequencies. The peculiarity of comb filter formants is that they are odd-numbered ($F_2 = 3F_1, F_3 = 5F_1$, etc.) Some vowels show similar patterns in their magnitude spectrum since the vocal tract is, in first approximation, a tube closed at one end that also favours odd-numbered resonant frequencies. This signifies that vowels and guitar tones are characterized by similar acoustic signatures, although the systems that produce them are structurally different. Previous attempts of locating formants within the instrument's body (the resonator) failed. From this, we learned that in order to establish perceptual analogies between vowel sounds and guitar sounds, it suffices to find similarities between the acoustical signatures of the sounds, regardless of their cause.

From speech perception to phonetics and singing pedagogy

While investigating verbal timbre descriptors commonly used by guitarists, we discovered that some of them refer to phonetic gestures: open, oval, round, thin, closed, nasal, hollow, etc. For example, when guitarists describe a guitar sound as round, it would signify that it sounds like a vowel produced with a round-shaped mouth, such as the vowel [ɔ]. In fact, the location of the comb filter formants along the frequency axis for a normal plucking position is similar to the location of the formants of a subset of vowels.

Linguists have defined distinctive features of speech such as openness, acuteness and laxness. For example, the vowel [i] is acute and tense; singers would qualify it as "pointed". The vowel [a] is open. The vowel [u] is closed; singers would describe it as "dark". Similar adjectives were used to qualify guitar tones which are perceived thinner (acute) and more nasal when plucked close to the bridge, and more closed and hollow when plucked close to the middle of the string; in the normal position – close to the tonehole – the guitar tones are perceived round and open. We noted a clear correspondence between the plucking position along the string, the frequency location of the induced comb filter formants and the association with certain vowels. The perceived nasality is explained by the broadening of the comb filter formants as the plucking gets closer to the bridge.

From phonetics to sonetics

In a listening experiment we conducted, listeners were asked to associate speech sounds to guitar tones. The choice of vowels was consistent with the qualifying adjectives. In their imitation of the guitar tones, the participants spontaneously chose different plosive consonants to emulate the different types of attack. This observation was the starting point for the development of systematic comparison between the elementary units of speech – the phonemes – and the elementary units of instrumental music, what we propose to call the *sonemes*. Phonemes and sonemes refer specifically to the timbral qualities of the sounds, regardless of pitch, duration and dynamics.

The aims of our questionnaire-based study were precisely the aims of a discipline B. Vecchione [163] calls *sonetics*: to establish relations between acoustical signals and characteristics of the signal producing gesture, and to identify regular associations between certain types of acoustical signals and perceptual dimensions of timbre.

Back to signal processing

Finally, we addressed the problem of the indirect acquisition of instrumental gesture parameters. Pursuing previous research on the estimation of the plucking position from a recording [48], we proposed a new method based on an iterative weighted least-square algorithm, starting from a first approximation derived from a variation of the autocorrelation function of the signal.

12.2 Applications and future directions

12.2.1 Control of sound synthesis

The results of this research may be applied to the control of sound synthesis. Though efficient sound synthesis algorithms exist – such as waveguide based physical models of plucked strings –, a sound synthesis algorithm serves little purpose when removed from the context of being played as an instrument, just as a note bears little meaning when removed from the context of a piece of music. The quality of the control parameters of a sound synthesis are vital in conveying the naturalness of the reproduction. A more thorough understanding of how performers control their acoustical instruments would better the development of digital instruments, equipping these with more meaningfully manipulable gestural interfaces.

In the course of this research, some elements of mapping between the interface and the produced sound have been clearly identified for the classical guitar. For example, the plucking position p is mapped to the delay D of the comb filter used as a plucking equalizer. The delay is expressed as the ratio of twice the absolute plucking position over the speed of sound on the string ($D = 2p/c$). The plucking angle may be mapped to the slope of a lowpass filter inserted in the string feedback loop. The exact correspondence is yet to be determined for this instrumental gesture parameter.

12.2.2 Talking guitars

Electrical guitarists have always attempted to convey a vocal quality with their guitar sounds. The "wah-wah effect" is the most familiar example. Other interesting effects can be obtained by enhancing the presence of formants in the guitar sounds. From the plucking point information, the comb filter formants could be localized and then thinned to obtain less nasal and more voice-like sounds.

12.2.3 New perceptual measures

As distinctive features of speech reveal themselves applicable to musical sounds, new measures can be developed on the basis of the frequency location of formant regions in the magnitude spectrum. These measures would be useful in the context of automatic timbre recognition and web-based search engines for sounds.

12.2.4 Exploring the timbre of other instruments

The interdisciplinary approach we propose for the study of the timbre of the classical guitar can be applied to other musical instruments, particularly stringed instruments, such as the violin, the viola and the cello. This would extend the development of sonetics which aims to exploring the relationship between instrumental gesture parameters (position, speed and force of the bow in the case of bowed-string instruments) and the perceptual dimensions of the produced timbre.

12.2.5 The musicology of the performer

Musicology is traditionally devoted to the historic study of composers by analyzing their work. It studies the art of composing through the written account of the compositional process – the score.

Very few musicologists study performers, most likely since a performance process is not tangible. This neglects the most fundamental aspect of musical creation, since this emerges at the level of the sound. In what and how Segovia and Rostropovitch were exceptionally gifted performers are questions that remain momentarily unanswered. By furthering the investigation of the correspondence between instrumental gesture, produced sound and perceived timbre, the art of performing will be better understood.

12.2.6 Pedagogical applications

In the context of teaching an instrument, the findings of sonetical research can contribute to the development of sophisticated pedagogical methods, enabling teachers to more efficiently communicate their art by promoting what Schneider calls *tone awareness*: "If the guitarist is aware of each of the timbral parameters that define the tone and is able to relate these parameters to the mechanical processes of the instrument and to his own actions, the player can change colours at will rather than by chance." [30]

Fig. 12.1 Symbolic picture illustrating a finger technique for the Ch'in, an ancient Chinese seven-string lute (from a Japanese manuscript copy of the *Yang-ch'un-t'ang-ch'in-pu*). 'The wild goose carrying a reed stalk in its bill' suggests to pluck a string with two fingers at the same time [25].

Appendix A

Autocorrelation

A.1 Autocorrelation function of an harmonic signal

The Fourier series form of a general periodic signal is

$$\begin{aligned} x(t) &= A_o + \sum_{n=1}^{\infty} A_n \cos(\omega_o n t) + B_n \sin(\omega_o n t) \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} X_n e^{j\omega_o n t} \end{aligned}$$

where

$$\begin{aligned} X_n &= \pi(A_n - jB_n) && \text{for } n > 0, \\ &= \pi(A_n + jB_n) && \text{for } n < 0, \\ &= 2\pi A_o && \text{for } n = 0. \end{aligned}$$

By definition, the infinite-duration autocorrelation function of a signal $x(t)$ is

$$a(\tau) = \lim_{T_D \to \infty} \frac{1}{2T_D} \int_{-T_D}^{T_D} x(t)x(t+\tau)dt$$

Replacing $x(t)$ by the expression of its Fourier series form, including a phase factor for

the shifted version $x(t+\tau)$, the autocorrelation becomes

$$\begin{aligned} a(\tau) &= \lim_{T_D\to\infty} \frac{1}{2T_D} \int_{-T_D}^{T_D} \left(\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} X_n e^{j\omega_o n t} \right) \left(\frac{1}{2\pi} \sum_{n'=-\infty}^{\infty} X'_n e^{j\omega_o n'(t+\tau)} \right) dt \\ &= \frac{1}{4\pi^2} \sum_{n=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} X_n X'_n e^{j\omega_o n'\tau} \lim_{T_D\to\infty} \frac{1}{2T_D} \int_{-T_D}^{T_D} e^{j\omega_o (n+n')t} dt \end{aligned}$$

where the factor $\lim_{T_D\to\infty} \frac{1}{2T_D} \int_{-T_D}^{T_D} e^{j\omega_o(n+n')t}dt$ equals 1 when $n+n'=0$ and 0 when $n+n'$ is a nonzero interger. As a result, $a(\tau)$ is nonzero only when $n'=-n$, which reduces the double sum to a single sum, leading to

$$a(\tau) = \frac{1}{4\pi^2} \sum_{n=-\infty}^{\infty} X_n X_{-n} e^{j\omega_o n\tau}$$

As the signal $x(t)$ is real, its transform is hermitian and therefore $X_n X_{-n} = X_n X_n^* = |X_n|^2 = |X_{-n}|^2$. The autocorrelation formula becomes

$$\begin{aligned} a(\tau) &= \frac{1}{4\pi^2} \sum_{n=-\infty}^{\infty} |X_n|^2 e^{j\omega_o n\tau} \\ &= \frac{1}{4\pi^2} \left(|X_o|^2 + \sum_{n=-\infty}^{-1} |X_n|^2 e^{j\omega_o n\tau} + \sum_{n=1}^{\infty} |X_n|^2 e^{j\omega_o n\tau} \right) \\ &= \frac{1}{4\pi^2} \left(|X_o|^2 + \sum_{n=1}^{\infty} |X_n|^2 2\cos(\omega_o n\tau) \right) \\ &= \frac{1}{4\pi^2} |X_o|^2 + \frac{1}{2\pi^2} \sum_{n=1}^{\infty} |X_n|^2 \cos(\omega_o n\tau) \\ &= A_o^2 + \frac{1}{2} \sum_{n=1}^{\infty} C_n^2 \cos(\omega_o n\tau) \end{aligned}$$

since

$$\begin{aligned} |X_o|^2 &= 4\pi^2 A_o^2 \\ |X_n|^2 &= \pi^2 (A_n^2 + B_n^2) = \pi^2 C_n^2 \end{aligned}$$

Therefore, the autocorrelation function of a periodic signal depends only on the Fourier coefficients, and not on the phases [58].

Appendix B

Symbols for Speech Sounds

B.1 Chart of tongue positions for vowels

On Fig. B.1, the vowels are placed according to tongue position (front/back, high/low).

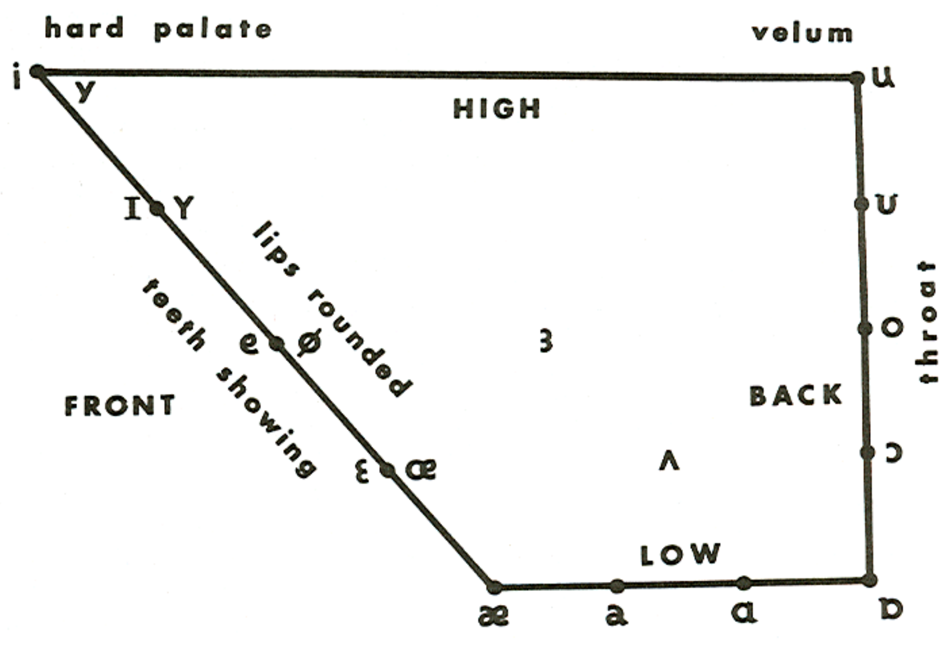


Fig. B.1 Chart of tongue positions for vowels. Vowels are indicated with International Phonetic Alphabet symbols [147].

B.2 IPA and *sound colour* symbols

Table B.2 gives the correspondence between the different symbols used to represented vowels as well as words in which the vowels are found (from [154] and [147]). The symbols are the common English or French spelling, the International Phonetic Alphabet symbol and the *sound color* notation as defined by Slawson, which is a two-letter convention that he believes was more evocative of most English speakers' phonetic intuitions.

IPA symbol	Sound color	English	Pronunciation (as in)
i	**ii**	*ee*	b*ee*t
ɪ	**ih**		bit
e	**ee**	*ay*	p*ay*
ɛ	**eh**	*eh*	pet
æ	**ae**		back
a	**aa**		bask
ɑ		*ah*	calm
ɒ			hot
ɔ	**aw**	*aw*	b*aw*
ɘ	**ne**		the
Λ	**ah**		cut
o	**oo**	*oh*	tone
ʊ	**uh**		put
u	**uu**	*oo*	b*oo*t
ʏ			German ü lax
IPA symbol	Sound color	French	as in
y		*u*	vu (also German ü tense)
ø	**oe**	*eu*	f*eu* (tense)
œ		*eu*	p*eu*r (lax)
œ̃		*un*	br*un*
ɛ̃		*in*	v*in*
ɔ̃		*on*	b*on*
ɑ̃		*an*	bl*an*c

Table B.1 International Phonetic Alphabet (IPA) symbols for English and French vowels, together with Slawson's sound color symbols and pronunciations.

Appendix C

Guitar Timbre Questionnaire and Ethics Form

Étude du timbre de la guitare

Pour définir les caractéristiques du timbre des sons qu'ils ou elles produisent, les guitaristes utilisent une multitude de qualificatifs, évoquant des matières, des couleurs, des sensations (visuelles, gustatives, tactiles, ...). Ci-dessous, vous trouverez deux listes de qualificatifs, la première en francais et la seconde en anglais.

La tâche consiste à :

- choisir 10 adjectifs qui définissent une caractéristique de timbre importante selon vous, en français ou en anglais selon votre préférence (entourez-les dans la ou les liste(s));
- pour chaque adjectif, donner une définition intuitive de cet adjectif (comment ça "sonne", ce que le timbre évoque, ...);
- pour chaque adjectif, expliquer comment on obtient ce timbre (mode de jeu sur la guitare);
- pour chaque adjectif, proposer un contraire et un synonyme, ainsi qu'une traduction dans l'autre langue en précisant la justesse de la traduction (si possible).

Pour effectuer cette tâche, vous pouvez utiliser votre propre expérience mais aussi toute autre référence comme des ouvrages ou des entrevues avec des guitaristes professionnels avec qui vous pourriez être en contact. Dans ce cas, veuillez préciser vos sources.

Si vous vous sentez inspiré(e), vous pouvez bien sûr soumettre les descriptions de plus de 10 adjectifs. Aussi, si vous pensez à un qualificatif qui n'est pas cité dans la liste et que vous jugez approprié, vous pouvez le choisir et le définir à la place d'un de ceux des listes proposées. Commentaires et suggestions seront les bienvenus.

En français : sombre, brillant, lumineux, mince, épais, mouillé, sec, opaque, chocolaté, doux, sucré, velouté, pulpeux, juteux, crémeux, laiteux, transparent, duveteux, florissant, vitré, cassant, métallique, cuivré, fibreux, laineux, confus, mat, voilé, spongieux, creux, nasal, nasillard, ovale, naturel, plein, émoussé, chaleureux, résonnant, rond, incisif, ouvert, fermé, sourd, clair, dur, mou, amer, large, étroit, lisse, rugueux, ...

En anglais : dark, bright, tinny, thin, thick, wet, dry, opaque, chocolaty, fudgy (super-chocolaty), buttery, sweet, sugary, velvety, fleshy, juicy, creamy, milky, transparent, feathery, blossom, glassy, metallic, brassy, naily, edgy, crisp, fibrous, wooly, muddy, veiled, spongy, swimming, hollow, woody, nasal, oval, full, full bodied, dull, mellow, warm, resonant, round, sharp, open, closed, clear, hard, soft, sweet, bitter, broad, narrow, smooth, rough, ...

Appendix D

Definitions of Guitar Timbre Descriptors

Appendix E

Publications

E.1 Thesis

Caroline Traube, *Digital Signal Processing Techniques for Estimating the Plucking Point on Stringed Instruments*, Engineer degree thesis, Center for Computer Research in Music and Acoustics, Stanford University, 2000.

E.2 Peer-reviewed conference articles related to the thesis topic

- Caroline Traube and Julius O. Smith III, "Estimating the plucking point on a guitar string", in *Proc. Conference on Digital Audio Effects*, Verona, Italy, pp. 153-158, 2000.

- Caroline Traube and Julius O. Smith III, "Extracting the fingering and the plucking points on a guitar string from a recording", in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, pp. 7-10, 2001.

- Caroline Traube and Philippe Depalle. "Deriving the plucking point location along a guitar string from the least-square estimation of a comb filter delay", in *Proc. Canadian Conference on Electrical and Computer Engineering*, Montreal, Canada, 2003.

- Caroline Traube, Philippe Depalle and Marcelo Wanderley. "Indirect acquisition of instrumental gesture based on signal, physical and perceptual information", in

Proc. International Conference on New Interfaces for Musical Expression, Montréal, Canada, pp. 42-47, 2003.

- Caroline Traube and Philippe Depalle, "Extraction of the excitation point location on a string using weighted least-square estimation of a comb filter delay", in *Proc. Conference on Digital Audio Effects*, London, England (UK), pp. 188-191, 2003.

- Caroline Traube, Peter McCutcheon and Philippe Depalle. "Verbal descriptors for the timbre of the classical guitar", In *Proc. Conference on Interdisciplinary Musicology*, Graz, Autria, 2004.

- Caroline Traube and Philippe Depalle "Timbral analogies between vowels and plucked string tones", in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Montréal, Québec, Canada, 2004.

- Caroline Traube and Philippe Depalle. "Phonetic gestures underlying guitar timbre description", in *Proc. International Conference on Music Perception and Cognition*, Evanston (IL), USA, 2004.

E.3 Communications

- *Extracting acoustical, gestural and perceptual information from recorded guitar tones.* Communication présentée pendant la Semaine Canadienne d'Acoustique 2002 organisée par l'Association Canadienne d'Acoustique du 9 au 11 octobre 2002 à Charlottetown, Île-du-Prince-Edouard.

- *Towards the modeling of instrumental gesture: deriving mechanical, perceptual and gestural parameters from the signal analysis of recorded instrumental tones.* Graduate Colloquium, Faculty of music, McGill University, 8 November 2002.

E.4 Co-supervision of graduate students

- Nadia Lavoie (D. Mus. - flute)

- Olivier Bélanger (D. Mus. - electroacoustic composition)

- Jehan Julien Filatriau (engineer)
- Nicolas D'Alessandro (engineer)

References

Guitar acoustics

[1] A. H. Benade, *Fundamentals of Musical Acoustics*. London: Oxford University Press, 1976.

[2] M. Brooke, *Numerical Simulation of Guitar Radiation Fields using the Boundary Element Method*. Phd thesis, University of Wales, Cardiff, 1992.

[3] J.-l.-R. d'Alembert, "Investigation of the curve formed by a vibrating string," in *Acoustics*, pp. 119–130, Lindsay, 1747.

[4] C. Gough, "The theory of string resonances on musical instruments," *Acustica*, vol. 49, pp. 124–141, 1981.

[5] N. H. Fletcher, "Plucked strings - a review," *The Catgut Acoustical Society Newsletter*, no. 26, pp. 13–17, 1976.

[6] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. New York: Springer-Verlag, 2e ed., 1998.

[7] D. E. Hall, *Musical Acoustics*. Sacramento: Thomson learning, 2002.

[8] E. V. Jansson, "Analogies between bowed string instruments and the human voice: Source-filter models," Tech. Rep. QPSR3/1966, Speech Transmission Laboratory, Royal Institute of Technology (KTH), Massachusetts Institute of Technology, Cambridge, Massachussets, 1966.

[9] E. V. Jansson, "Coupling of string motions to top plate modes in a guitar," tech. rep., Speech Transmission Laboratory, Royal Institute of Technology (KTH), Stockholm, Sweden, 1973.

[10] E. V. Jansson, "Acoustics for the guitar player," in *Function, Construction, and Quality of the Guitar* (E. V. Jansson, ed.), pp. 7–26, Stockholm: Royal Swedish Academy of Music, 1983.

[11] K. A. Legge and N. Fletcher, “Nonlinear generation of missing modes on a vibrating string,” *Journal of the Acoustical Society of America*, vol. 76, pp. 5–12, July 1984.

[12] M. E. McIntyre, R. T. Schumacher, and J. Woodhouse, “On the oscillations of musical instruments,” *Journal of the Acoustical Society of America*, vol. 74, no. 5, pp. 1325–1345, 1983.

[13] D. C. Miller, *The Science of Musical Sounds.* New York: McMillan, 1916.

[14] P. Morse, *Vibration and Sound.* New York: McGraw-Hill, 1948.

[15] M. Pavlidou, *A Physical Model of the String-Finger Interaction on the Classical Guitar.* Phd thesis, University of Wales, Cardiff, 1997.

[16] A. D. Pierce, *Acoustics. An Introduction to its Physical Principles & Applications.* New York: McGraw-Hill Book Company, 1989.

[17] B. E. Richardson, *A Physical Investigation into Some Factors Affecting the Musical Performance of the Guitar.* Phd thesis, University of Wales, Cardiff, 1982.

[18] I. Sloane, *Classic Guitar Construction.* London: Omnibus Press, 1976.

[19] G. P. Walker, *Towards a Physical Model of the Guitar.* Phd thesis, University of Wales, Cardiff, 1991.

[20] H. A. K. Wright, *The Acoustics and Psychoacoustics of the Guitar.* Phd thesis, University of Wales, Cardiff, 1996.

Guitar playing techniques

[21] P. Bone, *The Guitar and Mandolin (2d ed.).* London: Schott and Co., 1972.

[22] A. Company, *Las seis cuerdas.* Milano: Edizioni Suvini Zerboni, 1965.

[23] C. Duncan, *The Art of Classical Guitar Playing.* Princeton, New Jersey, USA: Summy-Birchard Music, 1980.

[24] J. W. Duarte, *The Bases of Classic Guitar Technique.* Kent: Novello publications, 1975.

[25] R. H. van Gulik, *The Lore of the Chinese Lute. An Essay in the Ideology of the Ch’in.* Tokyo: Sophia university Press, 1940.

[26] TV5, *An Interview with Alexandre Lagoya.* TV5 television station.

[27] E. Pujol, *The Dilemma of Timbre on the Guitar*. Buenos Aires: Ricordi Americana, 1960.

[28] P. Roch, *A Modern Method for the Guitar: School of Tarrega*. New York: Schirmer, 1921.

[29] G. C. Santisteban, *Carcassi Method for Classic Guitar*. USA: Ashley Publications Inc., 1967.

[30] J. Schneider, *The Contemporary Guitar*. Berkeley and Los Angeles, California: University of California Press, 1985.

[31] F. Sor, *Method for the Guitar, trans. A. Merrick, ed. F. Harrison*. London: Robert Cooks, 1896.

[32] J. Taylor, *Tone Production on the Classical Guitar*. London: Musical New Services Ltd., 1978.

Guitar tone analysis and synthesis

[33] K. Bradley, M.-H. Cheng, and V. L. Stonick, "Automated analysis and computationally efficient synthesis of acoustic guitar strings and body," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1995.

[34] G. Cuzzucoli and V. Lombardo, "Physical model of the plucking process in the classical guitar," in *Proc. International Computer Music Conference*, (San Francisco, USA), pp. 172–179, 1997.

[35] G. Cuzzucoli and V. Lombardo, "A physical model of the classical guitar, including the player's touch," *Computer Music Journal*, vol. 23, no. 2, pp. 52–69, 1999.

[36] C. Erkut, V. Välimäki, M. Karjalainen, and M. Laurson, "Extraction of physical and expressive parameters for model-based sound synthesis of the classical guitar," in *AES108th Convention*, (Paris), 2000.

[37] J. Hiipakka, *Implementation and Control of Real-Time Guitar Synthesizer*. Master thesis, Helsinski University of Technology, 1999.

[38] D. A. Jaffe and J. O. Smith, "Extensions of the karplus-strong plucked-string algorithm," *Computer Music Journal*, vol. 7, no. 2, pp. 56–69, 1983.

[39] M. Karjalainen, V. Valimaki, and Z. Janosy, "Towards high-quality sound synthesis of the guitar and string instruments," in *Proc. International Computer Music Conference*, pp. 56–63, 1993.

[40] M. Karjalainen, V. Valimaki, and T. Tolonen, "Plucked-string models: From the karplus-strong algorithm to digital waveguides and beyond," *Computer Music Journal*, vol. 22, no. 3, pp. 17–32, Fall 1998.

[41] M. Karjalainen, H. Penttinen, and V. Valimaki, "Acoustic sound from the electric guitar using dsp techniques," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 773 –776, 2000.

[42] K. Karplus and A. Strong, "Digital synthesis of plucked-string and drum timbres," *Computer Music Journal*, vol. 7, no. 2, pp. 43–55, Summer 1983.

[43] H. Penttinen and V. Valimaki, "A time-domain approach to estimating the plucking point of guitar tones obtained with an under-saddle pickup," in *Applied Acoustics.*, vol. 65, no. 12, pp. 1207–1220, Dec. 2004.

[44] D. Radicioni, L. Anselma, and V. Lombardo, "A segmentation-based prototype to compute string instruments fingering," in *Proc. of the Conference on Interdisciplinary Musicology (CIM04)*, (Graz, Austria), 2004.

[45] J. O. Smith, "Efficient synthesis of stringed musical instruments," in *Proc. International Computer Music Conference*, (Tokyo), pp. 64–71, 1993.

[46] J. O. Smith III, *Physical Audio Signal Processing: Digital Waveguide Modeling of Musical Instruments and Audio Effects.* Center for Computer Research in Music and Acoustics (CCRMA), Stanford University: Web published at http://ccrma.stanford.edu/ jos/pasp/, 2004.

[47] T. Tolonen and V. Välimäki, "Automated parameter extraction for plucked string synthesis," in *International Symposium on Musical Acoustics*, (Edinburgh, Scotland), pp. 245–250, 1997.

[48] C. Traube, *Digital Signal Processing Techniques for Estimating the Plucking Point on Stringed Instruments.* Engineer degree thesis, Center for Computer Research in Music and Acoustics, Stanford University, 2000.

[49] C. Traube and J. O. Smith, "Estimating the plucking point on a guitar string," in *Proc. Conference on Digital Audio Effects*, (Verona, Italy), pp. 153–158, 2000.

[50] C. Traube and J. O. Smith, "Extracting the fingering and the plucking points on a guitar string from a recording," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (Mohonk Mountain House, New Paltz, New York), pp. 7–10, 2001.

[51] C. Traube, "Spectrum estimation using pisarenko harmonic decomposition. application to guitar tones.," tech. rep., Electrical and Computer Engineering Department, McGill University, Montreal, Quebec, Canada, 2002.

[52] C. Traube and P. Depalle, "Deriving the plucking point location along a guitar string from a least-square estimation of a comb filter delay," in *Proc. Canadian Conference on Electrical and Computer Engineering*, (Montreal, Quebec, Canada), May 2003.

[53] C. Traube, P. Depalle and M. Wanderley, "Indirect acquisition of instrumental gesture based on signal, physical and perceptual information," in *Conference on New Interfaces for Musical Expression (NIME-03)*, (McGill University, Montreal, Quebec, Canada), pp. 42–47, May 22-24 2003.

[54] C. Traube and P. Depalle, "Extraction of the excitation point location on a string using weighted least-square estimation of a comb filter delay," in *Proc. Conference on Digital Audio Effects*, (London, UK), pp. 188–191, Sept. 2003.

[55] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Janosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis," *Journal of the Audio Engineering Society*, vol. 44, pp. 331–353, May 1996.

Digital signal processing

[56] J. C. Brown, "Computer identification of musical instruments using pattern recognition with cepstral coefficients as features," *Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1933–1941, 1999.

[57] T. Galas and X. Rodet, "An improved cepstral method for deconvolution of source-filter systems with discrete spectra: Application to musical sound signals," in *Proc. International Computer Music Conference*, (Glasgow, UK), pp. 82–84, 1990.

[58] W. M. Hartmann, *Signals, Sound, and Sensation*. Springer-Verlag New York Inc.: AIP Press, 1998.

[59] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: John Wiley & Sons, Inc., 1996.

[60] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, 1996.

[61] C. L. Lawson and R. J. Hanson, *Solving Least-Square Problems*. Prentice Hall Inc., 1974.

[62] M. Pluckette and C. Lippe, "Getting the acoustic parameters from a live performance," in *Proc. International Conference on Music Perception and Cognition*, pp. 328–333, 1994.

[63] T. F. Quatieri, *Discrete-time Speech Signal Processing. Principles and Practice.* Engelwood Cliffs, New Jersey: Prentice-Hall, 2002.

[64] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 1, pp. 24–33, 1977.

[65] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals.* Engelwood Cliffs, New Jersey: Prentice-Hall, 1978.

[66] E. D. Scheirer, *Extracting Expressive Performance Information from Recorded Music.* Master of science in media arts and sciences, Massachusetts Institute of Technology, 1995.

Gesture modelling

[67] C. Cadoz, "Instrumental gesture and musical composition," in *Proc. International Computer Music Conference*, (Köln, Germany), pp. 1–12, 1988.

[68] C. Cadoz and M. M. Wanderley, "Gesture - music," in *Trends in Gestural Control of Music* (M. M. Wanderley and M. Battier, eds.), pp. 71–94, Paris, France: Ircam - Centre Pompidou, 2000.

[69] F. Delalande, "La gestique de Gould: éléments pour une sémiologie du geste musical.," in *Glenn Gould, Pluriel* (G. Guertin, ed.), pp. 83–111, Louise Courteau Editrice Inc., 1988.

[70] P. Depalle, S. Tassart, and M. M. Wanderley, "Instruments virtuels," *Résonance*, September 1997.

[71] A. Gabrielsson, "The performance of music," in *The Psychology of Music* (D. Deutsch, ed.), pp. 501–602, San Diego, CA: Academic Press, 1999.

[72] H. Heijink and R. G. J. Meulenbroek, "On the complexity of classical guitar playing: Functional adaptations to task constraints.," *Journal of Motor Behaviour*, vol. 34, no. 4, pp. 339–351, 2002.

[73] D. Martino, "Notation in general - articulation in particular (1966)," in *Perspectives on Notation and Performance* (B. Boretz and E. Cone, eds.), pp. 102–113, New York: W. W. Norton and Co., 1976.

[74] M. M. Wanderley and P. Depalle, “Interfaces homme-machine et création musicale,” in *Contrôle gestuel de la synthèse sonore*, pp. 145–163, Hermes Science Publications, 1999.

[75] M. M. Wanderley, *Performer-Instrument Interaction: Applications to Gestural Control of Sound Synthesis*. PhD thesis, Université Pierre et Marie Curie - Paris VI, 2001.

[76] M. M. Wanderley and P. Depalle, “Gestural control of sound synthesis,” *Proc. IEEE*, vol. 92, no. 4, pp. 632–644, 2004.

Modelling and automatic recognition of timbre

[77] E. B. Egozy, *Deriving Musical Control Features from a Real-Time Timbre Analysis of the Clarinet*. Master thesis, Massachussetts Institute of Technology, 1995.

[78] K. Jensen, “The timbre model - discrimination and expression,” in *Proceedings of MOSART Workshop on Current Research Directions in Computer Music*, (Barcelona, Spain), 2001.

[79] K. Jensen, *Timbre Models of Musical Sounds*. PhD thesis, University of Copenhagen, 1999.

[80] T. Machover, “Hyperinstruments - a progress report 1987-1991,” tech. rep., MIT Media Laboratory Massachussets Institute of Technology, 1992.

[81] K. D. Martin, E. D. Scheirer, and B. L. Vercoe, “Music content analysis through models of audition,” in *ACM Multimedia - Workshop on Content Processing of Music for Multimedia Applications*, (Bristol, UK), 1998.

[82] K. D. Martin, “Toward automatic sound source recognition: Identifying musical instruments,” July 1-12 1998.

[83] N. Orio, “The timbre space of the classical guitar and its relationship with the plucking techniques,” in *Proc. International Computer Music Conference*, (Beijing, China), pp. 391–394, 1999.

[84] I. Fujinaga, “Machine recognition of timbre using steady-state tone of acoustic musical instruments,” in *Proc. International Computer Music Conference*, 1998.

Timbre perception

[85] ASA, *Acoustical Terminology.* New York: American Standard Association, 1960.

[86] W. W. Gaver, “How do we hear in the world? explorations in ecological acoustics,” *Ecological Psychology*, vol. 5, pp. 285–313, 1993.

[87] J. Grey, “Multidimensional perceptual scaling of musical timbres,” *Journal of the Acoustical Society of America*, vol. 61, no. 5, pp. 1270–77, 1977.

[88] J. M. Hajda, R. A. Kendall, E. C. Carterette, and M. L. Harshberger, “Methodological issues on timbre research,” in *Perception and Cognition of Music* (I. Deliège and J. Sloboda, eds.), pp. 253–306, Hove, UK: Psychology Press, 1996.

[89] S. Handel, *Listening: an Introduction to the Perception of Auditory Events.* Cambridge, MA: MIT Press, 1990.

[90] H. L. V. Helmholtz, *On the Sensation of Tone, Fourth German edition (1877), translated by Alexander J. Ellis (1885).* New York: Dover Publications Inc., 1954.

[91] E. Hermann-Goldap, “Ueber die klangfarbe einiger orchesterinstrumente,” *Annalen der Physik*, vol. 28, pp. 979–985, 1907.

[92] R. Plomp and H. J. M. Steeneken, “Effect of phase on the timbre of complex tones,” *Journal of the Acoustical Society of America*, vol. 46, p. 409, 1969.

[93] R. Plomp, *Aspects of Tone Sensation.* New York, New York: Academic Press, 1976.

[94] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff, “Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes,” *Psychological Research*, vol. 58, pp. 177–192, 1995.

[95] J.-C. Risset and D. L. Wessel, “Exploration of timbre by analysis and synthesis,” in *The Psychology of Music* (D. Deutsch, ed.), pp. 113–169, San Diego La Jolla, CA: Academic Press Series in Cognition and Perception, 1999.

[96] J. F. Schouten, “The perception of timbre,” *Reports of the Sixth International Congress on Acoustics*, vol. GP-6-2, 1968.

[97] C. E. Seashore, *Psychology of Music (original work published in 1938).* New York: Dover, 1967.

[98] G. von Bismarck, “Sharpness as an attribute of the timbre of steady sounds,” *Acustica*, vol. 30, pp. 159–72, 1974.

Timbre semantics

[99] A. C. Disley and D. M. Howard, "Timbral semantics and the pipe organ," in *Proc. Stockholm Music Acoustics Conference*, (Stockholm, Sweden), 2003.

[100] S. Donnadieu, *Représentation mentale du timbre des sons complexes et effets de contexte*. PhD thesis, Université René Descartes - Paris V Sciences Humaines - Sorbonne, 1995.

[101] A. Faure, *Des sons aux mots, comment parle-t-on du timbre musical ?* PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, 2000.

[102] R. Fitzgerald, *Performer-dependent dimensions of timbre: identifying acoustic cues for oboe tone discrimination.* PhD thesis, The University of Leeds, School of Music, 2003.

[103] R. A. Kendall and E. C. Carterette, "Verbal attributes of simultaneous wind instrument timbres. II - adjectives induced from Piston's *Orchestration*," *Music Perception*, vol. 10, pp. 469–502, 1992.

[104] J. S. Kerrick, D. C. Nagel, and R. L. Bennet, "Multiple ratings of sound stimuli," *Journal of the Acoustical Society of America*, vol. 45, no. 4, pp. 1014–1017, 1969.

[105] N. Lavoie, *Notions Fondamentales d'Acoustique et de Psychoacoustique comme aide à l'Exploration du Jeu Flûtistique*. Unpublished thesis. D. Mus., Faculté de musique, Université de Montréal, 2003.

[106] W. L. Martens and C. N. W. Giragama, "Relating multilingual semantic scales to a common timbre space," in *AES 113th Convention*, (Los Angeles, USA), October 2002.

[107] W. L. Martens, C. A. Marasinghe, C. N. W. Giragama, and A. P. Madurapperuma, "Topic-dependent adjective use in japanese and sinhala: Selection of adjectives differentiating guitar sounds," in *Proc. Seventh International Workshop on Human Interface Technology*, (Aizu-Wakamatsu, Japan), pp. 27–34, November 2000.

[108] H. Meinl, "Regarding the sound quality of violins," *Journal of the Acoustical Society of America*, vol. 29, pp. 817–822, 1957.

[109] A. Melka, J. Štěpánek, and Z. Otčenášek, "Czech and german verbal description of violin sound properties: Multidimensional analyses of survey data," *ACUSTICA - acta acustica*, vol. 82, no. Suppl. 1, p. 214, 1996.

[110] O. Moravec and J. Štěpánek, "Verbal description of musical sound timbre in czech language," in *Proc. Stockholm Music Acoustics Conference*, (Stockholm Sweden), 2003.

[111] W. Piston, *Orchestration.* New York: W. W. Norton & Co., 2000.

[112] L. N. Solomon, "Semantic approach to the perception of complex sounds," *Journal of the Acoustical Society of America*, vol. 30, pp. 421–425, 1957.

[113] J. Štěpánek and Z. Otčenášek, "Listener common and group perceptual dimensions in violin timbre," in *Proc. Stockholm Music Acoustics Conference*, (Stockholm, Sweden), August 2003.

[114] E. Samoylenko, S. McAdams, and V. Nosulenko, "Systematic analysis of verbalisations produced in comparing musical timbres," *International Journal of Psychology*, vol. 31, no. 6, pp. 255–278, 1996.

[115] G. von Bismarck, "Timbre of steady sounds: a factorial investigation of its verbal attributes," *Acustica*, vol. 30, pp. 146–159, 1974.

[116] C. Traube, P. McCutcheon, and P. Depalle, "Verbal descriptors for the timbre of the classical guitar," in *Conference on Interdisciplinary Musicology (CIM'04)*, (Graz, Austria), April 15-18 2004.

[117] C. Traube and P. Depalle, "Timbral analogies between vowels and plucked string tones," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Montréal, Québec, Canada), May 17-21 2004.

[118] C. Traube and P. Depalle, "Phonetic gestures underlying guitar timbre description," in *Proc. International Conference on Music Perception and Cognition*, (Evanston (IL), USA), August 2004.

Linguistics

[119] N. Chomsky and M. Halle, *The Sound Pattern of English.* New York: Harper and Row, 1968.

[120] J. Clarck and C. Yallop, *An Introduction to Phonetics and Phonology.* Wiltshire: Blackwell, 1990.

[121] R. Jakobson, C. G. M. Fant, and M. Halle, *Preliminaries to Speech Analysis. The distinctive Features and Their Correlates.* Massachusetts Institute of Technology, Cambridge, Massachussets: The MIT Press, 1967.

[122] R. Jakobson, *Child Language, Aphasia, and Phonological Universals.* The Hague: Mouton, 1968.

[123] R. D. Kent, *The Speech Sciences.* San Diego, London: Singular Publishing Group, Inc., 1997.

[124] S. R. Paget, *Human Speech.* New York, Harcourt: Brace & Co., 1930.

[125] K. L. Pike, *Phonemics: a Technique for Reducing Languages to Writing.* Ann Arbor: University of Michigan Press, 1947.

[126] G. O. Russel, *Speech and Voice.* New York: The Macmillian Company, 1931.

[127] E. W. Scripture, “The study of english speech sounds by new methods of phonetic investigation,” *Proceedings of the British Academy*, vol. 11, 1923.

Speech analysis and speech perception

[128] P. Delattre, A. M. Liberman, F. S. Cooper, and L. G. Gerstwan, “An experimental study of the acoustic determinants of vowel color; observations on one- and two-formant vowels synthesized from spectrographic patterns,” *Word*, no. 8, 1952.

[129] L. Eskenazi and D. G. Childers, “Acoustic correlates of vocal quality,” *Journal of Speech and Hearing Research*, vol. 33, pp. 298–306, 1990.

[130] M. P. Haggard, “Perception of semi-vowels and laterals (abstract),” *Journal of the Acoustical Society of America*, vol. 46, p. 115, 1969.

[131] S. Hawkins and K. Stevens, “Acoustic an perceptual correlates of the non-nasal/nasal distinction for vowels,” *Journal of the Acoustical Society of America*, vol. 77, no. 4, pp. 1560–75, 1985.

[132] D. Klatt, “Review of selected models of speech perception,” in *Lexical Representation and Process* (W. D. Marlsen-Wilson, ed.), Cambridge, Mass.: MIT Press, 1989.

[133] A. M. Liberman, F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy, “Perception of the speech code,” *Psychol. Rev.*, vol. 74, pp. 431–461, 1967.

[134] A. M. Liberman and I. G. Mattingly, “The motor theory of speech perception revised,” *Cognition*, no. 21, pp. 1–36, 1985.

[135] B. C. J. Moore, *An Introduction to the Psychology of Hearing.* Academic Press, 1997.

[136] G. E. Peterson and H. L. Barney, “Control methods used in the study of the vowels,” *Journal of the Acoustical Society of America*, vol. 24, 1952.

[137] L. C. W. Pols, "Perceptual space of vowel-like sounds and its correlation with frequency spectrum," p. 463, 1970.

[138] W. R. Tiffany, "Vowel recognition as a function of duration, frequency modulation and phonetic context," *J. Speech & Hearing Disorders*, vol. 28, 1953.

Singing voice

[139] G. Bjorklund, "Analysis of soprano voices," *Journal of the Acoustical Society of America*, vol. 33, pp. 575–582, 1961.

[140] G. Bloothooft and R. Plomp, "Spectral analysis of sung vowels. II. the effect of fundamental frequency on vowel spectra," *Journal of the Acoustical Society of America*, vol. 77, no. 4, pp. 1580–8, 1985.

[141] G. Bloothooft and R. Plomp, "Spectral analysis of sung vowels. III. characteristics of singers and modes of singing," *Journal of the Acoustical Society of America*, vol. 79, no. 3, pp. 852–64, 1986.

[142] R. Husson, *La voix chantée*. Paris: Gauthier-Villars. Collection Science et Techniques d'Aujourd'hui, 1960.

[143] N. Isshiki, "Regulatory mechanism of voice intensity variation," *J. Speech and Hearing Research*, vol. 7, 1964.

[144] L. Lehmann, *How to sing. Translation of "Meine Gesangkunst" by Richard Aldrich.* New York: The Macmillan Co., 1910.

[145] J. Sundberg, "Effects of the vibrato and the singing formant on pitch," *Musicologica Slovaca*, no. 6, pp. 51–69, 1978.

[146] J. Sundberg, "The perception of singing," in *The Psychology of Music* (D. Deutsch, ed.), pp. 171–214, San Diego La Jolla, CA: Academic Press Series in Cognition and Perception, 1999.

[147] W. Vennard, *Singing - the Mechanism and the Technic*. New York: Carl Fischer, 1967.

Music and language

[148] R. Alsop, *Using Aspects of Language in Computer Based Composition: Three Approaches to Current Australian Texts*. Master thesis, University of Melbourne, Australia, 1999.

[149] B. G. Levman, "The genesis of music and language," *Ethnomusicology*, vol. 36, Spring/Summer 1992.

[150] J.-J. Nattiez, *Music and Discourse. Toward a Semiology of Music. (Translator: Carolyn Abbate)*. New Jersey: Princeton University Press, 1990.

[151] A. D. Patel and J. R. Iversen, "Acoustic and perceptual comparison of speech and drum sounds in the north indian tabla tradition: An empirical study of sound symbolism," in *Proc. International Congress of Phonetic Sciences*, (Barcelona, Spain), pp. 328–333, 2003.

[152] R. Tsur, *What Makes Sound Patterns Expressive: The Poetic Mode of Speech-Perception*. Durham N, C.: Duke UP, 1992.

[153] R. Tsur, "Onomatopoeia: Cuckoo-language and tick-tocking the constraints of semiotic systems," *Web published at http://cogprints.ecs.soton.ac.uk/archive/*, 2001.

[154] W. Slawson, *Sound Color*. Berkeley and Los Angeles, California: University of California Press, 1985.

[155] G. P. Springer, *Language and Music: some Parallels and Divergencies. For Roman Jakobson*. The Hague: Mouton & Co., 1956.

[156] T. Wishart, *On Sonic Art*. York: Imagineering Press, 1985.

[157] J. Wolf, "Speech and music, acoustics and coding, and what music might be 'for'," in *Proc. International Conference on Music Perception and Cognition*, (Sydney, Australia), pp. 10–13, 2002.

[158] J. E. Youngblood, *Music and language: Some Related Analytical Techniques*. PhD thesis, Department of Music Theory, Indiana University, 1960.

Miscellaneous

[159] H. Cowell, *New Musical Resources*. New York: Knopf, 1930.

[160] E. M. von Hornbostel and C. Sachs, "Systematik der musikinstrumente. ein versuch," *Zeitschrift für Ethnologie*, vol. 46, no. 4-5, pp. 553–90, 1914.

[161] O. Laske, "An epistemic approach to musicology," in *Music processing* (G. H. (Ed.), ed.), Ephraim Nissan, JAI Press, 1991.

[162] B. Bell, *Acquisition et Représentation de Connaissances en Musique*. Phd thesis, Faculté des Sciences et Techniques de Saint-Jérôme Université de Droit, d'Économie et des Sciences d'Aix-Marseille, 1990.

[163] B. Vecchione, "Les sciences et les technologies de la musique – la révolution musicologique des années 1970-1980," in *Le Fait Musical – Sciences, Technologies, Pratiques, actes du Colloque "Musique et Assistance informatique"* (B. V. et B. Bel (Eds.), ed.), (Marseille (France)), pp. 69–128, 1990.